REPORT RESUMES

ED 020 148 SE 004 809

BIOGRAPHICAL INFORMATION AND THE PREDICTION OF MULTIPLE CRITERIA OF SUCCESS IN SCIENCE.

BY- TAYLOR, CALVIN W. AND OTHERS

NATIONAL AERONAUTICS AND SPACE ADMIN., WASH., D.C.

REPORT NUMBER NASW-105 PUB DATE FEB 66

EDRS PRICE MF-$0.50 HC-$5.00 123P.

DESCRIPTORS- *COLLEGE STUDENTS, *CREATIVITY, *SCIENTIFIC MANPOWER, *TALENT IDENTIFICATION, CAREERS, MULTIPLE CHOICE TESTS, PREDICTION, RESEARCH, RESEARCH METHODOLOGY, SECONDARY SCHOOL STUDENTS, NATIONAL AERONAUTICS AND SPACE ADMINISTRATION, CREATIVITY CHECK LIST, CREATIVITY RATING SCALE,

REPORTED ARE THE RESULTS OF SEVERAL RELATED STUDIES SUPPORTED BY THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA). THESE STUDIES ATTEMPTED TO (1) DETERMINE THOSE CHARACTERISTICS OF SCIENTISTS WHICH ARE VALID IN DIFFERENTIATING BETWEEN SCIENTISTS OF HIGH AND LOW LEVEL PRODUCTIVITY AND CREATIVITY, AND (2) UTILIZE THESE CHARACTERISTICS IN DEVELOPING AN EASILY ADMINISTERED AND SCORED BIOGRAPHICAL INVENTORY WHICH WOULD AID IN THE IDENTIFICATION OF SCIENTIFIC TALENT AT THE COLLEGE LEVEL. MUCH OF THE DOCUMENT IS DEVOTED TO SEVERAL STUDIES INVOLVING THE SEQUENTIAL DEVELOPMENT OF THE INVENTORY THROUGH FOUR FORMS. DATA AND A DISCUSSION OF THE RESULTS ARE PROVIDED FOR EACH FORM. STUDIES INVOLVING THE FIRST THREE FORMS ARE CONCURRENT VALIDITY STUDIES. THE FOURTH STUDY REPORTS HOW THE INSTRUMENT PERFORMED IN A FOLLOW-UP STUDY AND PROVIDES INFORMATION ON ITS VALIDITY IN A PREDICTIVE, LONGITUDINAL SITUATION. ALSO DISCUSSED ARE THE ITEM CONTENT OF THE INVENTORY AND OTHER RESEARCH ACTIVITIES DEALING WITH BIOGRAPHICAL INFORMATION. APPENDED ARE THE RATING SCALES FOR NASA SCIENTISTS. (DS)

ED020[illegible]

BIOGRAPHICAL INFORMATION AND THE
PREDICTION OF MULTIPLE
CRITERIA OF SUCCESS IN SCIENCE

by

CALVIN W. TAYLOR [illegible] ELLISON

[illegible]

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

SE004 809

BIOGRAPHICAL INFORMATION AND THE PREDICTION OF MULTIPLE CRITERIA OF SUCCESS IN SCIENCE

By

Calvin W. Taylor, Robert L. Ellison
and Michael F. Tucker

September, 1965

SUPPORTED BY: National Aeronautics and Space Administration
Research Project NASw-105

MONITORS: Robert Lacklen, Allen Gamble, and Ray Metcalf

Published by

The Creativity Research Institute
of The
Richardson Foundation
As Submitted Without Editorial Change

February, 1966

ACKNOWLEDGMENT

The time and sincere cooperation of the scientists and the personnel directors at each of the NASA research centers is greatly appreciated for without their generous help this study could not have been undertaken.

TABLE OF CONTENTS

Page

LIST OF TABLES. iv
LIST OF FIGURES vi

Chapter

1 INTRODUCTION 1
2 THE FORM A STUDY OF THE BIOGRAPHICAL INVENTORY 5
3 THE FORM B STUDY OF THE BIOGRAPHICAL INVENTORY 26
4 THE FORM C STUDY OF THE BIOGRAPHICAL INVENTORY 47
5 THE FORM C-1 STUDY OF THE BIOGRAPHICAL INVENTORY 57
6 EXAMPLES OF ITEM CONTENT OF THE BIOGRAPHICAL INVENTORY . 92
7 OTHER RESEARCH ACTIVITIES WITH BIOGRAPHICAL INFORMATION. 96
REFERENCES. 105
APPENDIX. 109

LIST OF TABLES

Table Page

1 Validities, Intercorrelations, Number of Items Per Biographical Score, Means and Standard Deviations for Sample 1 (179 NASA Scientists). 16

2 Validities, Intercorrelations, Number of Items Per Biographical Score, Means and Standard Deviations for Sample II (178 NASA Scientists). 17

3 Criterion Intercorrelations, Validities, Means and Standard Deviations for Form B Sample 1 (N = 148 NASA Scientists) 31

4 Criterion Intercorrelations, Validities, Means and Standard Deviations for Form B Sample 1 (N = 152 NASA Scientists). 37

5 Criterion Means, Standard Deviations, Intercorrelations and Cross Validities for Sample One in the Form C Study of the Biographical Inventory (N = 390). 49

6 Criterion Means, Standard Deviations, Intercorrelations and Cross Validities for Sample Two in the Form C Study of the Biographical Inventory (N = 379). 50

7 Organizational Breakdown for the Form C Study. 53

8 Average Cross Validities for Appropriate Keys Across the Two Odd - Even Sub-samples for each Organizational Group and Each Criterion in the Form C Study of the Biographical Inventory 54

9 Criterion Means, Standard Deviations, Intercorrelations and Cross Validities for the Total Sample (Sample 1) in the Form C-1 Study (N = 622 NASA Scientists). 63

10 Maximum Multiple Prediction of the Criteria in the Sample 1 Analysis . 66

11 Criterion Intercorrelations, Key Score Intercorrelations, Means, Standard Deviations, and Cross Validities for Form C-1 Sample II (N = 269). 69

12 Maximum Multiple and Partial Prediction of the Criteria in the Sample 11 Analysis 71

13 Criterion Intercorrelations, Key Score Intercorrelations, Means, Standard Deviations, and Cross Validities for Form C-1 Sample III (N = 56) 72

14 Criterion Intercorrelations, Key Score Intercorrelations, Means, Standard Deviations, and Cross Validities for Form C-1 Sample IV (N = 57) 74

Table Page

15 Maximum Multiple and Partial Prediction of the Criteria in the Sample IV Analysis. 76

16 Criterion Intercorrelations, Key Score Intercorrelations, Means, Standard Deviations, and Cross Validities for Form C-1 Sample V (N = 269) 77

17 Criterion Intercorrelations, Key Score Intercorrelations, Means, Standard Deviations, and Cross Validities for Form C-1 Sample VI (N = 149) 79

18 Maximum Multiple Prediction of the Criteria in the Sample VI Analysis . 82

19 Criterion Intercorrelations, Cross Validities, Means and Standard Deviations for Form C-1 Item Analysis - Odd Sample (N = 308). 86

20 Criterion Intercorrelations, Cross Validities, Means and Standard Deviations for Form C-1 Item Analysis - Even Sample (N = 314) 88

LIST OF FIGURES

Figure Page

1 The Sections of the Biographical Information Inventory (Form A), Number of Items per Section, and a Description of the Item Content per Section. 6

Chapter 1

Introduction

Since the launching of the Sputniks and the final awakening of the nation to the fact that Russia is producing approximately twice as many scientists and engineers as the U. S. whose technical training is apparently not inferior to our own (Golovin, 1963), the importance of identifying creative scientific talent has received a great deal of national publicity. However, the need for creative scientific talent has been recognized by some psychologists and manpower specialists years earlier. As early as 1955, research had been accomplished by enough psychologists so that the first of the Utah Creativity Conference series could be held on the Identification of Creative Scientific Talent with National Science Foundation support, and since that time an increasing number of researchers have undertaken studies on creativity (Taylor, 1956, 1958, 1959, 1961, 1962, and 1963). The principal investigator of the present project and Robert Lacklen, personnel director of NASA, as early as 1956 had discussed that each had decided independently that biographical information would probably be the best single means of identifying creative scientific talent. Intermittent discussions were carried on until 1959 when the present project was initiated. Thus the present study had a long history before it was started. This report presents the results of a project supported by the National Aeronautics and Space Administration (NASA) which was designed to do something to improve the quality of scientific talent available to the space agency and to the nation as a whole--in some ways a small but nevertheless an important step toward the solution of a crucially important problem.

When this study was initiated in 1959, biographical information was considered to be one of the most promising means of identifying creative scientific talent. Previous research from a variety of investigators (D. Taylor, 1957; Roe, 1951, 1963; Sprecher, 1957, etc.) had indicated that biographical information was a potentially promising technique for the identification of creative scientific talent, although no one had made a definitive attempt to exploit this avenue of approach. The approach had, however, demonstrated its usefulness in a variety of other settings for predictive purposes, e.g., for identifying successful salesmen (Kornhauser and McMurry, 1941), for predicting academic success (Fittner, 1945; Sorenson, 1950; Hansen, 1950) for identifying leadership ability in the Army after World War II (Adjutant General's Office, 1946), and others.

The intent in this present study was to exploit the biographical approach and thus to determine and more fully understand the experiences, backgrounds, opinions, self-images, attitudes, etc., which would aid in differentiating between the highly productive and/or creative scientists and those who were less productive and/or creative. In essence, the study attempted to determine what types of life history situations--self-perceptions, attitudes, etc.--tended to be more common to the more successful scientists and to ascertain the results of these situations as they were personified in the individual scientists.

When the biographical characteristics, experiences, self-descriptions, etc., were identified which would differentiate the more successful scientists from those who were less successful, the practical goal was to utilize these characteristics in developing an easily administered and scored biographical inventory which would aid in the identification of scientific talent at the college level. Hopefully, the inventory could be rewritten for the early high school level so that it could be used as a vocational guidance instrument, and thereby, high school students who had scientific potential could be encouraged to further their development.

Prior to the initiation of this large-scale NASA-supported research project, two studies were under way at the University of Utah to explore the relationship of biographical information to scientific accomplishment. The first was completed by Ellison (1960, 1962) in which 527 biographical items were administered to 71 advanced graduate students in the physical sciences. A large number of items was used to explore more adequately the potential of this type of instrument. Empirical keys were constructed to predict ratings of creativity, productivity, and the general characteristics of a successful scientist. An item alternative analysis was carried out and the resulting initial validities were very high, ranging from .91 to .94. No cross validation was attempted in view of the small sample size. Based on the results of this study a new form of the Biographical Inventory was constructed and administered as part of a very intensive study on 107 Air Force Scientists by C. W. Taylor, Smith, Ghiselin, and Ellison (1961). Empirical keys were developed via an item alternative analysis on the basis of the following six criteria: judged work output, supervisory ratings on overall performance, productivity in written work, supervisory rating of drive-resourcefulness, originality of written work, and status-seeking "organization-man" tendencies. As before, the initial validities obtained were very high for all six criteria and again cross validation was not attempted in view of the small sample size, but the best items from both this study and the one by Ellison were identified and retained for future use in the NASA project. In this study of Air Force scientists various a priori biographical keys proved to be among the best predictor scores used in the study. These keys were developed on the basis of the study by Ellison and also on the basis of the current state of knowledge of creativity research at that time.

In both of these studies, the items that were keyed and retained for use in future research were somewhat arbitrarily selected. In other words, they were not identified strictly in terms of the usual statistical significance requirements. This was done with the conviction that a consistent relationship even in the low levels of validity across studies and samples was a better method of item selection in the long run than a statistically significant correlation in any one study. All items so identified across a series of studies would probably meet the requirements of statistical significance as the sample size increased. The typical correlations between the item alternative and the criterion on the items retained for future research tended to be rather low, ranging from .20 to .40 with a rather small percentage (but at least 5%) of the sample choosing each alternative. The items so selected to form a longer combined test resulted in the high initial validities, even though each item alternative accounted for only a small percentage of the variance. We have sometimes described the Biographical Inventory with its many items and alternatives as an instrument consisting of a great many little oars, with each oar pulling only slightly in the right direction, but with all the oars in concert a powerful pull is exerted. The approach has some actuarial features in that experience tables have been constructed with information about each item so that the valid information is utilized to the fullest possible extent.

Chapter 2

The Form A Study of the Biographical Inventory

Based primarily on the two prior studies, a new form (Form A) of the Biographical Inventory (hereafter called the BI) was constructed for administration to NASA scientists. The instrument, as it was administered, consisted of 300 multiple choice items with the majority of the items containing either 4 or 5 alternatives. The three hundred items were subjectively classified into four sections. Additional information about the Inventory can be gained by examining Figure 1 which shows the sections of Form A of the BI, a brief description of the item content in each section and the number of items per section. By inspecting Figure 1, it becomes readily apparent that the items were very heterogeneous in nature. Thus the BI is not restricted to a narrow definition of biographical experiences. From our perspective the biographical approach should attempt to measure not only previous life history experiences (including past environmental effects on a person), but also assess the outcome or manifestation of the heredity-environment combination as it is personified in the individuals studied. ✓Thus, almost any trait, life history situation, experience, or self-description, etc., that was thought to have some relevance to the problem was considered if it could be cast into multiple choice form. The four and five alternative multiple choice format was utilized for a number of reasons: (a) it was more adaptable to describe life history situations and experiences; (b) it was more acceptable to the scientists; (c) it permitted a more detailed analysis, which could be reduced to a lesser number of alternatives if the item analysis results so indicated; (d) it provided an opportunity to see if any non-linear relationships existed between the alternatives of an item and the criteria. If there were any non-linear characteristics, they could be scored

Figure I

The Sections of the Biographical Information Inventory (Form A)

Number of Items per Section, and a Description of the Item

Content per Section.

Section Number	Sections of the BI	No. of Items per Section	Brief Description of Item Content per Section
1	Developmental History	75	Factual information about childhood activities, experiences, characteristics of the environment, etc., as well as subjective evaluations and descriptions of what was important to the subject, sources of derived satisfaction, dissatisfaction, etc.
2	Parents and Family Life	60	Descriptions both of factual and subjective nature of the subject's parents - their activities, interests, characteristics, patterns of discipline and encouragement, achievements, education, etc.
3	Academic Background	45	Academic experiences, attitudes, interests, achievements, study habits, etc. Length of time and types of subject matter studied.
4	Adult Life and Interests	120	Descriptions of leisure activities, interests, value preferences, self-descriptions and evaluations, achievements, job involvement, descriptions of what the subject would do in different hypothetical situations, work habits, etc.

accordingly to maximize the relationship of the item to the criteria. This point will be dsicsused in greater detail in a later report.

Description of the Criteria. In the selection and development of the criterion measures the intent was to use all available information at each research center visited and to collect a few well-constructed performance measures of our own for research purposes only. At the NASA research center where Form A was administered, an official rating score was available for each scientist included in the sample. This score was used by the research center to assess the overall performance of the men. The rating was made by the men's immediate supervisor and was reviewed by a higher level supervisor. This was the only criterion score which was based on the evaluation of two supervisors.

Since the dominant emphasis in this study has been the identification of creative scientific talent, a special effort was made to develop a criterion form to assess creative performance. Lacklen (1958) and Ghiselin (1963) had independently formulated similar definitions for measuring the creativity of a contribution, i.e., in Ghiselin's formulation "the degree to which a contribution restructures our universe of understanding" was the measure of a creative contribution and in Lacklen's formulation "the extent of the area which each contribution underlies," was the standard of measure. These conceptualizations were utilized as the basic rationale in the construction initially of a Creativity Check List, and later of a seven-step Creativity Rating Form. The Creativity Check List was combined on the same sheet with a Productivity Check List. Both check lists were adapted from the work of D. Taylor (1958) who developed the scales using the Thurstone technique for attitude measurement. In the case of the Creativity Check List, modifications were made in accordance with the work of Lacklen and Ghiselin as mentioned previously. The criterion form consisted of a series of statements with each statement having a predetermined numerical scoring value. However, these scoring values were unknown to the rater or supervisor. If the statement applied to or described the man being rated, it was checked by the rater, otherwise it was left blank. The man's rating on each type of performance, i.e., creativity and productivity, was the median scoring value of all of the statements that were checked.[1]

The sum of each man's publications and patents was also used as a criterion score. These two criterion scores were combined because at that time a computer program was not available which could handle more than three criteria at a time. This score was obtained from the scientists, hence, it must be interpreted with some caution since some distortion could have occurred. A recent study by Miltzer and Slatzer (1963) bears upon this point. In their study the correlation between scientists' reports of the number of their publications and an actual count of the number of their publications was found to be .51. These authors used this correlation to justify the use of reports rather than counts. This moderate relationship is certainly less than what might be desired but it is high enough to indicate that reports can be used with reservations if counts are not available. As stated before, some caution should be observed in the interpretation of these self reports.

The G. S. level of each scientist has also been used as a criterion measure with the hope that to some degree it reflects the achievement of the scientists but in the Form A study, once again computer program limitations did not allow the G. S. level to be used for item analysis purposes. It, of course, has the disadvantage that the G. S. level is partially a function of experience and it may also reflect other qualities of the scientist which are not necessarily related to merit, as was reported by Taylor et al. (1963).

---

[1] Copies of these and other criterion instruments used in these studies may be obtained as long as the supplies last by writing the authors.

The Creativity Rating Scale, which was designed to measure the same characteristics as the Creativity Check List, was administered three months after the Creativity Check List so that an estimate of the reliability of the two creativity measures could be obtained. On this form the rater checked the one descriptive statement which best described the man being rated. This form was filled out by the same raters as those who completed the Creativity Check List.

The Administration of the Biographical Inventory. The administration of the BI was carried out smoothly and efficiently, thanks to the great cooperation of all the scientists and other personnel at the research center. All of the higher level supervisor personnel of each research division participating in the study met with the psychologist from the University of Utah[2] who visited the research center and with other officials concerned. The nature of the study was discussed, together with the importance of obtaining accurate criterion ratings. The criterion forms and the Biographical Inventories were distributed at these meetings with instructions that they were to be returned in a weeks time. All of the ratings obtained were kept strictly confidential, i.e., they were used for research purposes only and no one at the research center had access to them. Code numbers were used on the Biographical Inventories to insure that the individual scientists would remain anonymous with only the visiting psychologist having access to the name for each code number. An information sheet was also distributed with each inventory explaining the nature of the study to the individual scientists. Out of the 455 Biographical Inventories which were distributed, 357 were completed and returned, 11 were returned with incomplete information, and 49 of the scientists were on leave. Thus, there was a balance of only 34 scientists who did not respond within approximately a week to the request for their cooperation. This is an excellent example of the high degree of cooperation which has been characteristic throughout the entire research project.

---

[2] Robert Ellison was the field psychologist who visited the NASA research centers. Gary Cooley also participated in one of the field trips.

The educational background of the scientists was as follows: 24% had a bachelor's degree, 38% had some graduate training but no graduate degree, 13% had a master's degree, 16% has some graduate work beyond the master degree, and 9% has Ph.D. degrees. The particular areas of scientific research and engineering under investigation at this research center dealt with problems of propulsion for both atmospheric and space flight. Also under investigation were chemical, electrical, and nuclear propulsion systems, as well as more exotic schemes such as those involving thermal and solar conversion processes.

The Data Analysis. A double cross-validation study was carried out as follows: the total sample of 357 scientists was numbered in consecutive order. All of the odd-numbered scientists were arbitrarily assigned to Sample I and all of the even-numbered scientists were assigned to Sample II. The sample size was 179 for Sample I and 178 for Sample II.

An item analysis was then carried out for each sample independently. Due to computer limitations at that time, only three criteria could be used in the item analysis for each sample. For Sample I these were: the Offical Rating Criterion, the Creativity Check List, and the Publications-and-Patents Criterion. For Sample II the criteria used in the item analysis were: the Official Rating Criterion, the Creativity Check List, and the Productivity Check List. In the item analysis, biserial correlations were utilized to determine the relationship of every alternative of each item to each of the three criteria. Since the Biographical Inventory had 300 questions and most of the questions contained five alternatives, there were approximately 1500 alternatives in the Biographical Inventory. The relationship of each of these alternatives to each of the three criteria in each sample was analyzed independently. Thus, approximately 9000 biserial correlations were computed. All of the computations were carried out on a Burroughs 205 medium speed computer.

For scoring purposes an alternative was scored if at least 5% of the sample chose the alternative and if the correlation of the alternative with the criterion was .20 or greater, regardless of sign. This level was selected on the basis of convenience and tradition. In both of the studies which preceded this Form A study, this percentage value and correlation level were used with the highly promising results mentioned previously. In terms of statistical significance, this does not quite meet the requirements for the .05 level of significance. The standard error for a biserial correlation where the correlation is assumed to be 0, and 5% of the sample respond to the alternative with a sample size of 179 is .118 so that a correlation of .23 would be significant at the .05 level. If more than 5% of the sample chose an alternative, then a lower correlation would be significant at this level.

If a correlation was plus .20 or greater, a weight of +1 was used, and a correlation of minus .20 or greater was scored -1. In this manner separate positive and negative scoring keys were constructed for each criterion. The literature did not reveal any clear superiority for a more complicated weighting system (for example, see Guilford, 1950, pp. 537-542). The sectional breakdown of the BI was also retained in the scoring keys; that is, for each criterion, the scores were as follows: Section 1--Developmental History up to Age 21, Section 2--Parents and Family Life, Section 3--Academic Background, Section 4--Adult Life and Interest, and a Total Score which was the algebraic sum of these four section scores. The empirical keys that were developed on the basis of Sample I were then applied to Sample II and vice versa so that a double cross-validation design was carried out. In order to determine the magnitude of the initial validities, the Official Rating Key which was constructed on the basis of Sample I was also applied to Sample I.

Seven a priori keys were also included in the data analysis. These keys were based on two studies mentioned previously (Ellison, 1960; Taylor et al., 1961) and upon the basis of clinical judgment. The best items from these two previous studies were given to two clinical consultants, together with definitions of seven traits thought to be important for creative performances in science. Their task was to classify each item into the one most appropriate category that they thought the item measured. All items on which there was no agreement between the two judges were not retained for this analysis. The seven traits were as follows: Professional Self-confidence, Low Sociability, High Self-sufficiency, Dedication to Work, Liking to Think, Intellectual Thoroughness, and Inner Directedness. These keys were applied to both Sample I and Sample II.

After the item analysis on both Sample I and Sample II was completed and the keys were constructed for each sample, two additional sets of keys were developed. These were a combined Official Rating Key (with subscores for Sections 1, 3, 4, plus the Total Score) and a combined Creativity Check List Key (with subscores for Sections 1, 3, 4, plus the Total Score). These keys were composed of items which had an average correlation of .20 or greater, with similar signs across the two samples and to which at least 5% of the total sample (Sample I plus Sample II) responded. Thus, these keys, when applied to both Sample I and Sample II, were indicative of the initial validity of the instrument if an item analysis had been carried out on the full sample of 357 scientists.

Two additional variables were also included in the Sample I analysis. The first of these was the total number of times each scientist was scored on the Sample Two Official Rating Key regardless of sign. Although no definitive evidence is available, it is suspected that this score reflects, at least to some degree, the independence of the subjects completing the Biographical Inventory, since the items with the low percentage values (i.e., the percentage of the sample which chose each alternative) are the ones which tend to be scored on the inventory. Thus, those scientists who tended to mark the alternatives which were chosen less frequently by the other scientists, as a whole, would tend to be higher on this score. The other additional variable included in the analysis was termed an Off-line Measure. This score was an exploratory attempt to develop a correction score for those scientists who tended to be classified in an incorrect manner by their biographical score; that is, the scientists could be classified in terms of those who either had a high biographical score in relation to their criterion score, those who had similar biographical and criterion scores, and those who had a low biographical score in relation to their criterion score. As seen from another perspective, this score measured how much error was made on each scientist in predicting his criterion score from his biographical score. Briefly, the procedure used in developing this score was as follows: on Sample I the Official Rating Criterion Score and the Sample II Official Rating Key Total Score were both equated in terms of means and standard deviations. Then the Key Score was subtracted from the Official Rating Criterion Score and a large enough constant was added to make all scores positive. In this manner a new distribution was created which described the subjects according to the type and the degree of the error in the initial prediction. This score was constructed for later use in an item analysis and it was included in this analysis to ascertain its characteristics. The plan was that the key derived from it would then be cross validated on an independent sample with the hope that it would aid in identifying an additional portion of the criterion variance, i.e., the false positives and the false negatives.[3]

All the correlations between the variables for each sample were computed on the IBM 7090 at Western Data Processing Center.[4] The validities,

---

[3] The reader may recognize that this score bears some relationship to the standard error of estimate:

$$\sigma_{yx} = \sqrt{\frac{\Sigma (Y-Y')^2}{N}}$$

where: Y=Measured value of a case we are trying to predict
Y'=Predicted value for each case
N=The total number of cases predicted

[4] The authors wish to acknowledge the facilities and the cooperation that was made available at Western Data Processing Center, University of California, Los Angeles.

intercorrelations, means, standard deviations, and number of items per biographical score for Sample I are shown in Table 1 and for Sample II in Table 2.

Results of the Form A Study. The Sample I criterion intercorrelations are shown for variables 1 through 7 in Table 1. Of particular interest are the correlations of the Creativity Check List, variable 2, with the Creativity Rating, variable 6, in Sample I. It will be remembered from the previous discussion that the Creativity Rating Form was administered three months after the Creativity Check List, and that the forms were radically different in nature; thus, this correlation can be taken as a conservative estimate of the reliability of the Creativity Criterion measures. This correlation was .69. It was definitely lower in Sample II due to the fact that complete data for the Creativity Rating were not available at the time this analysis was carried out and mean scores were used as an estimate where missing data were present.

In Table 1, the Sample I Official Rating Key scores (Sections 1 through 4 and the total) which were applied to Sample I criterion measures are variables 8 through 12. They give the initial validity for these keys on this sample for the Official Rating Criterion. They range from .53 for Section 2 to .79 for the total score against this criterion.

Variables 13 through 30 in Table 1 were constructed on the basis of Sample II and thus, the correlations of these scores with the criteria were all cross validity coefficients. From pilot studies it was determined that the Parents and Family Life Section, Section 2 of the Biographical Inventory, consistently did not hold up on cross validation. Thus, in these scores a total score was computed with and without Section 2 for each key. Table 1 shows that Section 2 of each of the three sets of

Table I

Validities, Intercorrelations, Number of Items per Biographical Score, Means and Standard Deviation for Sample I
(179 NASA Scientists)

	Score	No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	Number of Items Keyed	Mean	S.D.
Criteria	Official Rating	1	—																																													-	57.10	14.15
	Creativity Check List	2	65	—																																												-	43.64	12.81
	Productivity Check List	3	68	68	—																																											-	46.64	9.06
	Publications & Patents	4	52	46	37	—																																										-	6.28	6.55
	G. S. Level	5	44	25	37	55	—																																									-	12.27	2.21
	Creativity Rating	6	69	69	39	43	34	—																																								-	40.34	14.10
	Combined Creativity (2 + 6)	7	73	92	69	48	32	92																																								-	49.99	9.07
Sample I Keys	Official Rating Key, Section 1	8	70	54	53	32	38	58	60	—																																						32	40.55	4.58
	Section 2	9	53	37	38	30	30	49	47	50	—																																					19	40.10	3.71
	Section 3	10	55	46	41	36	26	45	49	41	36	—																																				22	39.89	5.82
	Section 4	11	72	67	54	66	46	60	69	64	42	56	—																																			51	41.44	15.75
	Total	12	79	69	59	66	47	66	73	74	55	71	94																																			124	52.13	24.19
Sample II Keys	Official Rating Key, Section 1	13	17	20	23	09	11	22	23	22	07	13	27	27	—																																	23	28.34	3.08
	Section 2	14	-06	07	02	07	08	-00	03	-05	-07	01	11	05	22	—																																19	29.86	2.05
	Section 3	15	32	29	25	28	17	28	31	25	17	65	37	47	31	14	—																															20	32.01	2.87
	Section 4	16	53	55	40	57	45	48	57	48	27	46	83	78	33	11	37	—																														60	36.36	11.92
	Total with Section 2	17	53	56	43	52	43	50	58	47	25	53	82	79	49	26	49	92	—																													122	36.57	14.37
	Total without Section 2	18	55	56	44	53	43	52	59	50	27	54	83	80	49	13	50	94	99	—																												103	36.78	14.03
	Creativity Check List Key Section 1	19	16	18	14	07	05	21	21	25	12	19	28	30	54	13	25	27	41	40	—																											22	38.74	2.66
	Section 2	20	01	-01	-02	02	-01	-00	-01	08	13	07	08	12	06	27	17	07	17	13	20	—																										19	39.58	2.42
	Section 3	21	35	33	32	25	14	27	33	31	18	82	42	55	15	04	69	34	45	46	26	14	—																									17	41.53	3.87
	Section 4	22	51	54	39	51	37	47	55	44	23	48	83	76	33	11	34	90	90	90	28	04	37	—																								60	42.49	9.55
	Total with Section 2	23	51	53	39	47	33	48	55	49	27	63	81	81	40	17	52	83	91	90	50	29	64	90	—																							118	42.21	13.31
	Total without Section 2	24	52	56	41	49	35	49	57	49	25	64	82	81	41	13	51	85	91	91	48	12	63	92	98	—																						99	42.57	12.80
	Productivity Check List Key Section 1	25	09	06	02	00	-03	06	07	10	04	09	14	15	38	02	12	17	24	24	38	08	10	18	25	25	—																					20	38.70	1.84
	Section 2	26	-14	-09	-13	06	-02	-15	-13	-08	-32	-06	07	-02	04	39	05	14	19	13	12	34	-02	10	15	10	-04	—																				25	38.62	2.52
	Section 3	27	12	11	10	06	02	08	10	15	-01	21	16	18	05	10	42	12	22	21	21	17	36	13	28	25	14	09	—																			13	39.23	1.65
	Section 4	28	42	44	35	42	34	37	44	39	16	32	71	63	24	07	17	77	71	72	19	05	22	77	68	70	16	18	10	—																		52	42.66	8.19
	Total with Section 2	29	39	41	31	42	30	33	40	36	11	33	72	63	32	18	29	79	78	78	33	19	28	76	76	75	33	37	30	90	—																	110	39.60	9.57
	Total without Section 2	30	46	47	37	43	32	40	47	42	20	38	75	68	34	09	31	80	79	80	33	12	31	80	78	78	36	14	30	91	97																			85	40.89	8.97
Sample I & II Keys	Combined Official Rating Key Section 1	31	56	50	50	30	25	46	52	75	39	37	52	61	48	01	33	46	52	54	36	13	30	44	51	50	15	-04	12	38	38	43	—																	24	41.07	2.59
	Section 3	32	47	40	34	34	19	36	41	37	29	90	51	66	18	08	71	44	54	55	24	14	87	46	67	66	11	-03	22	27	30	34	38	—																18	40.51	4.25
	Section 4	33	64	63	51	61	45	57	65	57	37	54	93	89	32	10	37	87	87	88	27	09	43	88	85	86	15	06	15	76	75	78	52	52	—															41	41.16	11.74
	Total	34	69	67	55	60	42	60	68	65	42	70	92	94	35	10	52	85	89	90	33	14	60	85	89	90	17	04	19	70	70	74	64	71	94	—														83	42.71	15.90
	Combined Creative Chk. List Key Sec. 1	35	46	50	43	32	19	45	52	57	27	31	55	59	45	07	29	52	61	61	56	23	29	53	63	61	18	09	18	45	51	53	68	33	57	63	—													27	40.44	2.54
	Section 3	36	46	48	38	33	21	43	49	40	26	86	55	68	20	11	67	46	57	57	26	17	87	50	71	71	10	02	27	29	35	37	40	91	55	72	40	—												16	39.79	3.38
	Section 4	37	61	68	50	59	41	55	67	55	33	54	94	88	33	13	38	89	90	90	30	10	44	91	88	90	18	09	17	77	78	81	51	51	96	94	59	56	—											33	41.24	11.94
	Total	38	64	70	55	57	39	58	70	60	35	65	91	91	36	13	48	85	89	90	36	15	57	87	91	92	19	08	22	71	73	77	59	64	93	96	69	70	97											76	41.39	15.74
Personality Keys	Professional Self Confidence	39	44	47	33	40	30	38	46	36	30	59	70	71	23	16	47	66	71	71	27	19	53	70	75	75	12	07	16	48	51	53	37	65	71	77	44	66	75	78	—									17	5.44	2.87
	Low Sociability	40	09	09	12	07	29	20	15	08	20	05	08	11	10	-43	07	06	08	10	03	02	06	07	08	08	-12	-25	-16	-03	-12	-06	13	09	09	11	07	10	05	08	15	—								14	4.03	2.31
	High Self Sufficiency	41	16	19	13	12	05	27	25	15	21	16	10	17	19	-09	16	05	14	15	10	04	20	10	17	17	03	-32	-08	-12	-13	-05	21	27	11	19	16	27	11	18	15	42	—							15	4.13	2.31
	Inner Directedness	42	07	12	05	08	02	12	13	16	12	05	19	17	12	-00	-01	21	19	20	10	-11	05	20	17	19	-01	-15	-06	11	05	10	12	05	24	20	17	11	22	22	16	07	11	—						10	26.31	12.17
	Dedication to Work	43	07	00	04	09	18	13	07	06	-07	05	10	07	03	14	10	14	19	17	12	16	11	13	18	16	-01	22	10	10	15	10	02	09	11	12	07	14	10	11	11	20	00	08	—					13	5.36	1.87
	Liking to Think	44	11	19	09	20	17	15	18	10	22	30	24	37	20	18	22	30	34	33	12	08	28	29	33	33	-03	08	-10	12	10	09	19	37	39	34	22	30	29	32	29	16	21	15	-00	—				14	4.32	1.88
	Intellectual Thoroughness	45	41	41	28	40	28	35	41	32	21	70	66	67	26	09	56	66	70	71	26	11	61	67	73	74	13	18	16	52	55	55	37	69	67	75	40	68	68	72	66	05	17	15	15	38	—			10	39.81	19 64
	# of Times Scored Official Rating Key	46	33	20	21	29	20	25	25	36	35	37	30	40	11	03	39	25	31	32	14	15	38	23	32	31	03	-11	04	-02	02	06	35	42	26	37	20	38	24	30	40	34	18	02	10	16	29	—		-	44.47	6.21
	Off Line Measure	47	48	10	26	00	03	18	15	22	29	01	-11	00	-34	-20	-19	-42	-47	-46	-26	-12	-12	-41	-41	-41	-15	-29	-09	-31	-40	-35	03	-08	-25	-21	-16	-12	-30	-26	-28	-01	01	-14	-11	-24	-32	02		-	43.62	13.29

Note:--Decimal points omitted; $r_{05} = .15$, $r_{01} = .20$.

Table II

Validities, Intercorrelations, Number of Items per Biographical Score, Means, and Standard Deviations for Sample II
(178 NASA Scientists)

	Score	No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	Number of Items Keyed	Means	S.D.
Criteria	Official Rating	1	--																																								-	54.89	13.71
	Creat. Check List	2	63	--																																							-	41.98	12.23
	Prod. Check List	3	66	62	--																																						-	44.76	09.54
	Publications & Patents	4	40	29	24	--																																					-	06.67	06.76
	G. S. Level	5	52	32	30	69	--																																				-	12.15	01.88
	Creativity Rating	6	55	57	54	26	26	--																																			-	40.20	09.63
	Combined Creativity	7	64	86	62	31	33	81	--																																		-	50.14	08.63
Sample I Keys	Off. Rating Key	8	31	23	13	31	34	25	26	--																																	32	41.47	03.90
	Section 2	9	06	12	-07	08	14	00	06	25	--																																19	40.81	03.02
	Section 3	10	29	34	14	15	15	22	30	15	21	--																															22	39.92	05.54
	Section 4	11	55	50	40	50	49	42	48	26	18	41	--																														51	40.92	12.19
	Total with Section 2	12	54	49	31	52	52	39	46	45	38	62	90	--																													124	43.51	17.18
	Total without Sec. 2	13	55	49	34	55	53	41	47	44	24	60	92	99	--																												105	42.72	16.32
	Creat. Ck. List Sec. 1	14	30	23	00	23	24	19	20	57	32	23	32	46	44	--																											39	41.26	04.06
	Section 2	15	03	00	-06	-02	01	-03	-01	13	31	06	14	16	12	23	--																										27	38.03	04.39
	Section 3	16	32	33	14	29	17	26	30	16	14	74	47	63	63	23	04	--																									25	39.76	04.60
	Section 4	17	50	42	34	42	40	36	41	20	16	39	92	85	87	30	17	45	--																								48	41.39	13.09
	Total with Section 2	18	53	48	29	42	40	39	45	36	31	54	88	91	91	54	31	63	91	--																							139	40.54	18.03
	Total without Sec. 2	19	54	48	32	45	42	40	46	34	24	55	89	92	93	50	19	64	93	98	--																						112	42.72	17.11
	Pub. and Pat. Sec. 2	20	25	17	-03	27	34	17	20	57	27	26	37	50	48	45	19	25	32	42	40	--																					30	40.54	03.44
	Section 2	21	-04	-01	-08	-02	-05	01	02	-01	29	09	-01	09	04	16	25	10	00	09	04	12	--																				34	42.99	03.17
	Section 3	22	17	13	02	11	10	12	15	06	21	56	24	40	38	17	06	46	23	33	33	19	23	--																			21	41.73	04.77
	Section 4	23	53	48	42	54	54	37	45	25	13	36	91	85	87	28	14	43	89	83	86	35	-03	24	--																		56	42.88	12.45
	Total with Section 2	24	49	45	33	50	53	38	45	33	25	52	85	89	90	39	22	54	83	86	86	54	19	44	92	--																	141	48.50	16.16
	Total without Sec. 2	25	51	46	35	52	54	39	46	34	19	52	87	90	91	37	16	56	85	85	87	54	05	41	93	98	--																107	45.24	16.03
Sample I & II	Comb. Off. Rating 1	26	41	32	17	26	26	32	33	52	25	19	31	42	41	61	12	26	25	40	36	36	24	15	26	32	34	--															24	41.06	03.01
	Section 3	27	40	34	20	28	25	26	32	16	22	81	45	64	63	25	05	76	46	59	60	26	03	58	47	60	60	13	--														18	40.56	04.21
	Section 4	28	60	50	41	50	49	40	48	24	17	39	93	86	88	33	14	44	94	88	90	33	-01	27	91	86	86	26	49	--													41	40.72	10.67
	Total Sections	29	64	53	40	51	50	44	52	33	25	57	89	92	93	44	14	61	88	91	93	40	05	41	87	89	90	44	68	94	--												83	42.69	14.27
	Comb. Creat. Key 1	30	36	47	21	27	25	31	42	52	22	22	36	45	45	66	16	28	32	49	47	38	15	11	34	41	41	67	18	35	45	--											27	40.32	02.68
	Section 3	31	39	48	25	24	20	36	44	13	16	75	53	65	65	28	09	78	53	67	68	21	06	50	51	60	60	23	83	54	69	30	--										16	40.06	03.50
	Section 4	32	55	54	43	47	43	40	49	21	16	40	93	85	87	30	18	45	94	88	90	30	05	26	91	85	86	28	46	95	90	35	53	--									33	40.74	11.17
	Total Sections	33	58	60	42	47	43	45	55	29	20	53	90	89	90	42	19	58	91	93	94	35	09	35	88	88	88	39	59	92	94	52	70	96	--								76	41.30	14.40
Personality Keys	Prof. Self-Conf.	34	40	48	31	35	31	35	44	06	17	52	69	69	70	23	12	51	72	72	73	13	02	39	68	68	69	17	63	72	75	28	69	72	77	--							17	05.24	02.76
	Low Sociability	35	06	01	-04	13	18	-03	00	15	21	03	09	14	11	17	05	-06	09	11	10	20	14	08	05	11	09	18	02	15	16	12	00	09	09	11	--						14	04.02	02.27
	High Self-Sufficiency	36	02	07	-10	03	10	12	11	26	24	04	16	20	17	32	21	06	16	24	20	28	17	10	08	20	16	20	04	18	20	19	08	13	15	19	43	--					15	42.36	22.17
	Inner Directedness	37	15	05	-07	00	03	06	06	06	19	08	12	15	13	16	09	08	11	17	16	16	14	17	11	16	15	16	03	13	15	13	08	11	13	11	24	22	--				10	25.90	11.77
	Dedication to Work	38	10	02	08	17	17	07	03	07	05	14	22	24	24	00	-02	11	27	21	23	05	-13	12	27	25	28	04	21	26	27	-03	17	24	23	28	17	02	-01	--			13	05.28	01.91
	Liking to Think	39	13	11	-03	05	-02	14	10	09	13	26	23	27	27	32	17	18	28	34	33	31	13	30	21	34	32	15	29	27	32	25	31	26	32	30	16	26	20	05	--		14	04.55	01.99
	Intellect. Thoroughness	40	38	39	27	33	31	36	42	12	10	58	63	66	68	22	14	51	65	66	68	23	-08	45	66	69	71	13	68	64	71	25	69	63	69	70	02	09	09	30	33	--	10	38.26	19.65

Note:--Decimal points omitted; $r_{05} = .15$, $r_{01} = .20$.

criterion keys consistently had either negative or essentially zero validities. Section 4, The Adult Life and Interests Section, consistently had the highest validity of any of the other sections, followed by Section 3 (the Academic Background Section), then Section 1 (the Developmental History Section), and finally Section 2 (the Parents and Family Life Section). The total score cross validities were .55 for the Official Rating Key score against the Official Rating Criterion; .56 for the scores from the Creativity Check List Key against the Creativity Check List Criterion and .37 for the Productivity Key against the Productivity Criterion.

The combined keys, that is, the keys that were based on both Sample I and Sample II, are variables 31 through 38 in Table 1. These are the keys that would be especially useful in further research, since they represent the best items from the analysis of both samples. These validities are somewhat higher than those obtained from the cross validated keys and not as high as the Sample I keys applied to Sample I, as one would expect, since this is not the same as a regular cross validation. As mentioned previously this was an estimate of the initial validity of the instrument if an item analysis had been performed on the total sample of 357 scientists.

The *a priori* personality keys were somewhat disappointing, as only four out of the seven keys correlated significantly at the .05 level with the criteria and two of these had only a low relationship though these were with the creativity criteria only. The Professional Self-confidence Key, which contained 17 items, proved to be the most valid *a priori* key with a validity coefficient of .47 with the Creativity Check List Criterion. The other *a priori* key which had substantial validity was the Intellectual

Thoroughness Key which correlated .41 with the Creativity Check List Criterion. The other two keys which produced significant criterion correlations were the High Self-sufficiency Key and the Liking to Think Key.

The Number of Times Scored on the Official Rating Key correlated .33 with the Official Rating Criterion and also correlated significantly with all of the other criteria, although not as high. This preliminary analysis of the Off-line Measure also proved intriguing since scores from this measure correlated .48 with the Official Rating Criterion and -.46 with the total score of the Official Rating Key; thus, it was hoped that when an item analysis was carried out against this Off-line Measure the biographical key derived from such an analysis would correlate positively with the criteria and negatively with the key, and thereby contribute substantially in a multiple correlation equation. A thorough discussion of the results of these attempts to develop correction scores will be presented in a later report.

The results for Sample II are similar to Sample I although the cross validities in general were not quite as high. The cross validity sufficient for the total score of the Official Rating Key against the Official Rating Criterion was .55. For the Creativity Check List Total Score Key the cross validity coefficient was .48 against the Creativity Check List Criterion, and the Publication and Patent Keys total score correlated .35 with the Publications and Patent Criterion.

The Official Rating Criterion was somewhat more predictable than the other criteria in the study, and the keys that were constructed on the basis of this criterion also gave the highest cross validities in the study. In Table 1 the correlation of variable 18, the Official

Rating Key Total Score without Section 2, with the combined Creativity Rating Criterion was .59, whereas the Creativity Check List Total Score without Section 2 correlated .57 against the same criterion. The stability of the scores from the Official Rating Keys and the predictability of the Official Rating Criterion is probably due to the fact that two raters completed these ratings, whereas in the other ratings only one supervisor was involved.

When one considers the magnitude of those cross validity coefficients that were obtained from only a single instrument, the Biographical Inventory, it is apparent that they represent a substantial degree of prediction. For comparative purposes, an illustration may be given. Psychologists have been working for some forty to fifty years on the prediction of grade-point average. A wide variety of tests, measuring a multitude of traits, both personality and intellectual, have been used in an attempt to predict this criterion. Yet, after some forty or fifty years of research, the typical validity coefficients that are reported in the literature for the prediction of grade-point average tend to range from .40 to .60. Usually, correlations of this magnitude are based on not just one instrument, but on two or three and perhaps several. Thus it can be seen that with just the Biographical Inventory a level of prediction against job performance criteria has been achieved that is very favorably compared with the results obtained from a problem that psychologists have been working upon for some forty or fifty years. Even more impressive evidence can be gained when the unreliability of the criterion is considered. Take, for example, the correlation of .69 between scores on the two creativity criteria mentioned previously that were obtained some three months apart. This correlation is probably a conservative estimate of the reliability of these scales although Guilford (1956a, p. 478) states that the reliability of ratings, even good ones, is about .60. But in any event, if the validity coefficient of a single biographical test score of .59 against the combined Creativity Criterion is corrected for unreliability in the criterion only, the cross-validity coefficient would increase to approximately .70. This is a very high level of prediction under any circumstances, especially for such an important external criterion as creativity.

The _a priori_ personality keys when applied to Sample II were not as valid as they were in Sample I, since only 2 of the variables had significant correlations with the criteria. These were the Professional Self-confidence Key with a validity coefficient of .48 against the Creativity Check List Criterion and the Intellectual Thoroughness Scale with a validity coefficient of .39 against the Creativity Check List Criterion. Scores from these two keys, which were designed to predict creative performance, were able to predict all the criteria at the .01 level; however, they had higher correlations with the creativity criteria than with the other criteria used in the study.

Discussion of the Results in the Form A Study. The generally high intercorrelations among the criteria were in marked contrast to the study by Taylor et al. (1961) on Air Force scientists. In Taylor's study, the criteria were obtained from a variety of sources yielding many different kinds of scores. In the present study, all of the criterion scores had their origin from either the scientist's immediate supervisor or from the scientists themselves, so that the number of sources utilized was considerably less than in Taylor's study. The Publication and Patents Criterion and the G. S. Level Criterion were, generally speaking, more independent of the other criteria used in this study. The relative independence of these criteria from the other criteria of success in science is a finding which is similar to the results found by Taylor although they were more related to each other than in the study by Taylor et al. In Taylor's study, multiple criterion measures were factor analyzed and a factor emerged which was identified as Productivity in Written Work, which was relatively independent and separate from a Quality of Written Work, and also from supervisors' ratings of performance on various traits. In Taylor's study, G. S. Level also emerged on a separate factor which was labeled Current Organizational Status. Again, this factor was independent from the majority of the other criteria used in the study, i.e., it contained no high loadings from the measures which defined the other factors of Quantity of Work, Originality of Written Work, etc.

As one might expect, the magnitude of the initial validity coefficient is partially a function of the sample size. In looking across the college study by Ellison (1960), the Air Force Study by Taylor et al. (1961), and the present study, there has been a consistent decline in the magnitude of the initial validities obtained as the sample sizes have increased. In the college study, with an N of 71, the highest initial validity was .94 and in the present study, the initial validity was .79 when the Sample I keys were applied back to Sample I, (N = 179). When the combined keys, that is, those that were built on both Sample I and Sample II, were applied to both subsamples, the average initial validity for the total score of the Combined Official Rating Key was .66. This was with a sample size of 354. Therefore, it seems likely that, at least with the present inventory, the level might be reached where the initial validity would be only slightly greater than average cross validity of the two subsamples, although the size of the sample where this would be true is as yet unknown. It is recognized that other variables also have a bearing on this question, such as the number of items and alternatives in the questionnaire, etc.

The disappointing performance of the Parents and Family Life Section of the BI is hard to explain. Certainly a variety of theoretical formulations as well as previous studies, partially those of Roe (1951, 1963), have indicated that the socialization procedures of the parents should have a substantial bearing on the outcome and development of the children. Indeed, other investigators have constructed instruments to measure these effects. Examples are the Parental Attitudes Research Inventory by Schaeffer and Bell (1958), and the questionnaire developed by Roe (1960), both of which were designed to specifically measure the effects of the parents and the family life upon the socialization of the children. In the present study there are a number of possible explanations for these results. Perhaps the present study failed to include appropriate items, or the items themselves were poorly constructed. Another reason that appears more probable is that when any one facet of parental behavior has been measured it does not provide enough information about how the other parental characteristics interact due to the complex network of interactions that exist between the subject's parents. Thus, by itself, the characteristic being measured appears unimportant. Perhaps the most likely explanation is that this area is extremely difficult to measure in any fashion if results are to be obtained that indicate that the socialization procedures have an influence on the performance of scientists in a research setting.

The relatively low cross validities obtained from scores on the Productivity Check List Keys against the Productivity Criterion were an unexpected phenomena. The highest cross validity obtained with one of these keys against the Productivity Check List Criterion was .37 as shown in Table 1. This was the Total Score without Section 2. These results may have been due to the implicit selection procedures of the investigators, as they selected items that were primarily related to the creativity aspects of the criteria. An alternate possibility is that the Productivity Check List Criterion, for some reason, was less reliable than the other criteria and hence less predictable. This possibility has some support from other sources. For example, in Sample I the Productivity Check List Criterion correlates slightly higher with the Official Rating criterion (.68) than does the Creativity Check List Criterion (.65). Yet, when the Official Rating Key for Sample I was applied to Sample I, it correlated higher with the Creativity Check List than with the Productivity Check List, i.e., the total score of the Official Rating Key, constructed on the basis of Sample I and applied to the Sample I criteria, correlated with the Creativity Check List Criterion .69 while the same key correlated .59 with the Productivity Check List Criterion. Thus, from another source, there was evidence that the Productivity Check List Criterion was less reliable and hence more difficult to predict from the items included in the Biographical Inventory.

As mentioned before, the validity coefficients of the *a priori* scores were rather disappointing. This was a preliminary attempt to determine the construct validity of some of the items in the inventory. What was really needed was a factor analysis of the best biographical items with landmark variables included as well as the criterion scores.

Such a study has recently been completed and the results will be presented in a later report.

A rather intriguing finding that came out of the analysis of the *a priori* scores concerns the relative magnitude of the validity coefficients of these scores predicting the Official Rating Criterion and the Creativity Check List Criterion. There is a noticeable tendency, especially in Sample I, but which is also discernible in Sample II, for the Creativity Criterion to be more predictable by these *a priori* scores than was the Official Rating Criterion. Yet, as mentioned previously, the Official Rating Criterion was overall the most predictable criterion in the study in terms of the empirical keys which were constructed. This evidently illustrates again the implicit techniques of the investigators in constructing and selecting items which were somewhat more related to the Creativity Criterion than the other criteria used in the study.

Chapter 3

The Form B Study of the Biographical Inventory

A new form of the Biographical Inventory was constructed (Form B) based on the best items of the previous administration. The "deadwood" items were eliminated from the previous form of the Biographical Inventory to make room for new items to be evaluated in this form with the hope that the percentage of "livewood" items in the new set of 300 items would be increased. The best items that were identified from Form A, that is, those items which worked consistently across both subsamples were subjectively reclassified into four substantive categories which appeared to be the most descriptive global categories of what the items were measuring. The criterion for the selection of an item into Form B required that each item have one or more alternatives with an average validity coefficient of .20 or greater with either the Creativity Criterion or the Official Rating Criterion across both of the subsamples with at least 5% of the total sample responding to the valid alternative. Other carry -over items included in Form B were those retained for control purposes such as the number of years of education, age, etc., and a few other items of interest from a theoretical perspective. The subscores into which the items were subjectively classified were as follows: (1) General Intellectuality, (2) Independence, (3) Professional Self-confidence, (4) Miscellaneous, and (5) the algebraic sum of the previous four scores. Two psychologists were responsible for this sorting of the items. In cases where there was disagreement as to the category in which the item should be placed, a discussion was carried out and, if disagreement still resulted, the item was classified into the miscellaneous category. These keys then represented the best accumulated information from the first study of the Biographical Inventory.

In the second study, a revised form of the Creativity Rating Scale was administered. It was the only criterion score collected for research purposes only. Available at the research center were a number of official evaluations and data were collected on the most appropriate of these. The selected ratings were on the following characteristics: Knowledge of Work, Initiative, Judgment, Industry, Reliability, and Cooperation. In addition, information was also collected from the scientists on the number of publications and patents that each had produced.

At the second NASA research center 370 Biographical Inventories were distributed and 354 were returned, leaving only 16 scientists who were not accounted for. Again this was considered to be an extremely satisfactory return. Once more the cooperation and interest of everyone concerned was very high and greatly appreciated. The 354 returned included nine who were on leave and four which were incompletely filled out. This left a maximum of 341 cases available for processing for which criterion data were available on our creativity rating. However, due to missing data on the criterion scores which were available at the research center, the final sample used in the data analysis was 300. Again the same procedure was followed in analyzing the data. Briefly, the total sample of 300 scientists was split in two subsamples of 148 and 152, as in the previous administration, and an item-alternative analysis was carried out for each sample in a double cross-validation design.

For the processing of these data a new computer program was utilized as well as a new computer. Arrangements were made at the Western Data Processing Center of the University of California at Los Angeles for the processing of the data. A biserial, point biserial item analysis program was written by Mr. Dan Morris, now of the University of California at

San Diego, specifically for analyzing the data. Since this program resulted in a substantial time saving, as well as providing additional research information, it will be described briefly. A maximum of 13 criteria may be used in the item analysis program. The computer first computes the mean and standard deviations of the criteria, then the relationship (either biserial or point biserial correlations or both) of each alternative of each question to each of the 13 criteria is analyzed individually. A standrad error for a biserial correlation is also computed for each alternative. The absolute number of scientists choosing each alternative is printed out as well as the percentage of the sample choosing each alternative. The program also computes the mean and standard deviations of the distribution of the scientists across the alternatives of each question. After completing these computations, the program computes the intercorrelations among the criteria and then scores a second sample of scientists on the basis of the item analysis performed on the first sample. The scores of each scientist for each of the criteria are then printed out. Several options are available on the type of print-out desired. For example, a separate score was obtained for each of the four sections of the Biographical Inventory, namely, the Developmental History Section, the Parents and Family Life Section, the Academic Background Section, and the Adult Life and Interests Section; finally a total score was computed based on the previous four. After these scores are computed, the cross validity coefficients are computed for each key against each of the criteria used in the analysis. After the program was "debugged," the computer performed all these operations, which included the calculation of approximately 36,000 biserial correlations, the means, and standard deviations of the questions and the criteria, scored both samples of scientists according to the magnitude of the correlations, and computed the cross validity coefficients and the criterion correlations in approximately 12 minutes. The use of computers has, needless to say, greatly facilitated this research. In fact, without computers such research could only be approximated. The intensive analyses that have been carried out have added greatly to the results that have been obtained.

The criterion intercorrelations, cross validities, and standard deviations for Ames Sample I are presented in Table 3. The first twelve variables were held up as criteria in an item analysis on the two samples; however, no results are shown from the item analysis of variable 12, the Form A Key total score. Variables 13 through 67 were the empirical keys which were constructed from an item analysis of Sample II and then applied to the scientists in Sample I. In other words, all of the correlations in this section are cross-validity coefficients. Variables 68 through 75 were various a priori predictor keys derived as follows: variable number 68, labeled Off-line Key #4, was one of the correction scores developed with the hope that it would help to identify those scientists with a large amount of error in their predicted criterion scores. Briefly, this score was obtained from an item analysis of Sample II where the criterion used was that portion of the criterion variance that had not been identified by the key from the first study. As mentioned previously, these correction scores will be discussed in a later report. Variable number 69, Rank in College, was an item which was taken from the Biographical Inventory and it read as follows: About what percentage of students did you surpass academically when you graduated from college? The alternatives were: A. 99%, B. 90%, C. 60%, D. 40%, E. Don't know. Variable 70 was a similar item but in this case the scientist described was high school performance.

Here the alternatives read: A. 99%, B. 90%, C. 80%, D. 60% and E. 50% or less. General Intellectuality, variable 71, was one of the keys which was developed from the first study of the Biographical Inventory as mentioned previously. This also applies to variable 72, Independence, variable 73, Professional Self-confidence, and variable 74, the Miscellaneous category of items. Variable 75, Independent Intellectuality, was strictly an a priori exploratory attempt to determine what relationship, if any, the learning of complex terms, symbols, and concepts on the students' own time would have to success in science. Examples of the terms to which the scientists responded whether they had learned them in school or on their own time were as follows: fugue, allegro, pollenation, volt, catalyst, spontaneous combustion, wolverine, electro magnet, AC and DC current, survival of the fittest, osmosis, and lepidoptera. If the scientists reported that they learned these terms prior to twelve years of age outside of school they were given two points for each one; if they had learned these terms or concepts between twelve and eighteen years of age outside of school they were given 1 point for each term to which this classification applied.

Results and Discussion of Sample I in the Form B Study. In Table 3, which presents the results from an analysis of Sample I, the first ten measures were all criterion scores and it will be noticed they were all characterized by moderately high interrelationships. The only exception was variable 3, Number of Patents, which was relatively independent of the other criteria used in the study. This variable, with a mean of .22 and a standard deviation of .58 was so skewed, with the overwhelming majority of the scientists having zero patents, that the relationships with other variables were to be expected. Scores from the Off-line Measure, variable

Table 3

CRITERION INTERCORRELATIONS, VALIDITIES, MEANS, AND STANDARD DEVIATIONS

FOR FORM B SAMPLE I (N = 148 N.A.S.A. SCIENTISTS)*

			1	2	3	4	5	6	7	8	9	10	11	12	Means	S. D.
Criteria	1. Creativity Rating		—												38.51	14.54
	2. Publications		46	—											5.46	5.69
	3. Patents		10	19	—										.22	.58
	4. G. S. Level		55	63	07	—									12.45	1.72
	5. Knowledge of Work		66	34	16	41	—								62.95	11.53
	6. Initiative		71	30	19	40	75	—							61.74	12.95
	7. Judgement		69	31	12	44	77	82	—						60.92	12.41
	8. Industry		34	11	-07	16	28	50	45	—					65.68	11.41
	9. Reliability		51	23	-05	31	51	62	66	67	—				67.08	11.15
	10. Cooperation		25	08	03	10	33	44	41	47	60	—			71.29	10.86
	11. Off-Line Measure		51	05	-18	18	29	39	39	20	42	24	—		50.45	12.34
	12. Form A Key Total Score		48	40	28	37	38	33	32	12	08	-03	-50	—	51.89	12.29
Sample II Empirical Keys	13. Section I	Creativity Key	05	09	-08	14	12	-00	-06	06	-04	-07	03	03	51.13	3.23
	14. Section II	"	13	00	-06	-10	03	16	11	17	19	15	11	02	53.11	3.02
	15. Section III	"	22	12	08	14	19	08	-09	03	02	-04	-23	46	49.02	2.66
	16. Section IV	"	49	54	22	55	34	35	35	16	11	03	-30	80	50.22	10.67
	17. Total Score	"	49	52	17	51	35	35	34	20	12	04	-29	80	53.48	12.34
	18. Section I	Publications Key	-03	19	10	20	01	-04	-05	-04	-05	-03	-17	13	52.02	3.11
	19. Section II	"	06	01	-03	06	-06	01	02	-12	17	12	-07	12	50.02	2.10
	20. Section III	"	15	26	11	21	17	16	06	-14	07	10	-20	34	49.95	3.23
	21. Section IV	"	42	56	23	49	29	28	29	-11	08	00	-36	79	46.60	9.96
	22. Total Score	"	37	55	23	49	27	26	23	13	09	03	-39	76	48.53	12.69
	23. Section I	Patents Key	-10	-01	07	-12	02	-02	-06	-07	-05	01	-12	00	51.54	2.00
	24. Section II	"	13	00	-13	-03	05	09	08	05	07	02	18	-04	52.73	1.64
	25. Section III	"	-08	-17	03	-24	09	06	-02	-11	-10	02	-01	-05	50.78	1.78
	26. Section IV	"	35	36	30	33	23	24	24	09	06	-01	-25	60	51.59	3.70
	27. Total Score	"	22	21	22	11	20	22	16	01	01	01	-18	39	56.52	5.12
	28. Section I	G. S. Level Key	-02	26	09	30	-04	-01	-03	-02	02	01	-06	02	51.17	2.39
	29. Section II	"	09	13	03	12	02	-01	03	02	07	06	-11	18	49.93	2.63
	30. Section III	"	21	27	05	29	21	13	07	07	-02	-04	-21	41	49.66	2.44
	31. Section IV	"	42	53	22	52	26	28	27	11	07	-01	-34	77	51.16	10.31
	32. Total Score	"	40	56	21	56	25	25	23	11	07	00	-35	75	51.92	13.58
	33. Section I	Know. of Work Key	03	02	04	09	12	02	02	-03	-13	-09	05	09	48.63	2.80
	34. Section II	"	10	20	02	10	11	10	05	03	10	03	-16	26	51.97	2.25
	35. Section III	"	35	21	11	35	28	16	16	01	-00	-06	-13	48	49.15	2.61
	36. Section IV	"	49	50	22	45	32	33	32	09	05	-04	-32	82	48.24	8.56
	37. Total Score	"	47	47	21	47	35	30	30	08	04	-06	-31	79	47.86	11.60
	38. Section I	Initiative Key	-03	-03	13	05	04	01	-02	-02	-04	-10	06	-08	51.01	2.95
	39. Section II	"	21	20	-09	08	14	13	10	03	15	09	10	12	51.62	2.25
	40. Section III	"	20	11	07	18	18	11	19	11	02	-11	-24	46	51.20	2.25
	41. Section IV	"	43	43	23	40	26	28	28	08	02	01	-31	88	49.04	7.05
	42. Total Score	"	43	40	22	39	29	28	28	09	05	-03	-32	77	52.67	9.83
	43. Section I	Judgment Key	04	03	-06	18	10	00	-01	-02	-06	-13	03	02	51.77	3.13
	44. Section II	"	15	14	04	09	13	17	11	07	19	10	01	14	52.25	2.09
	45. Section III	"	22	10	12	19	21	18	19	07	04	-04	-15	38	50.35	1.82
	46. Section IV	"	54	47	20	52	36	40	39	16	16	07	-21	76	50.82	5.70
	47. Total Score	"	52	44	17	54	40	38	37	15	18	02	-19	72	55.19	7.53
	48. Section I	Industry Key	-04	-04	06	-08	01	-04	-06	-01	-06	-07	02	-06	48.51	3.51
	49. Section II	"	20	14	-01	13	15	14	14	-03	03	01	08	14	51.83	1.60
	50. Section III	"	12	04	03	13	12	12	12	21	17	05	-17	30	50.17	1.66
	51. Section IV	"	39	33	20	38	23	29	30	11	08	-05	-34	73	47.37	5.41
	52. Total Score	"	34	26	17	30	23	24	23	11	06	-07	-27	67	47.80	7.06
	53. Section I	Reliability Key	-03	08	09	19	-06	-12	-16	-07	-14	-15	-02	-02	50.70	2.90
	54. Section II	"	14	05	-02	01	13	11	17	-07	08	05	06	08	51.24	2.00
	55. Section III	"	23	15	02	22	23	18	13	20	07	00	-19	41	47.64	2.87
	56. Section IV	"	43	58	22	54	31	35	30	12	11	02	-21	65	50.88	4.55
	57. Total Score	"	38	46	18	53	29	27	23	12	07	-06	-22	61	50.57	6.83
	58. Section I	Cooperation Key	02	03	07	13	07	-05	-05	-06	-04	-06	-05	08	51.50	2.70
	59. Section II	"	-02	10	02	-03	00	-03	-07	01	-05	-04	-08	05	50.46	1.96
	60. Section III	"	11	09	08	10	15	19	10	16	03	03	-11	25	49.17	2.43
	61. Section IV	"	25	24	-05	43	16	18	14	13	18	12	-04	28	51.75	3.21
	62. Total Score	"	18	21	04	31	18	16	08	13	07	05	-13	30	52.80	5.63
	63. Section I	Off-Line Key	-00	-12	-17	-03	-01	-04	-05	02	05	-01	31	-32	53.59	9.06
	64. Section II	"	-05	-16	-15	-17	01	03	04	-01	-03	01	17	-22	49.47	2.38
	65. Section III	"	-25	-18	-10	-13	-26	-20	-20	-12	-04	-04	33	-48	49.53	3.88
	66. Section IV	"	-32	-28	-24	-23	-20	-21	-18	-01	04	05	46	-40	48.98	11.89
	67. Total Score	"	-33	-27	-26	-18	-16	-17	-15	-02	04	03	50	-75	51.98	20.29
Other Predictors	68. Off-Line Key P4		21	16	-13	24	05	07	08	14	12	07	28	-07	54.70	8.45
	69. Rank in College (#167)		11	13	05	10	10	06	11	08	05	06	-14	21	73.83	15.36
	70. Rank in High School (#143)		15	01	-00	07	18	19	19	05	07	12	-10	26	82.35	13.78
	71. General Intellectuality		30	34	25	12	31	25	30	11	07	01	-39	71	51.69	4.60
	72. Independence		36	31	25	26	28	27	23	03	03	-08	-33	71	49.87	3.52
	73. Prof. Self-Confidence		47	46	24	42	29	28	26	13	09	-03	-37	65	49.78	6.17
	74. Miscellaneous Items		34	23	12	30	23	13	14	03	-03	04	-40	56	50.89	2.34
	75. Independent Intellectuality		-10	04	11	-02	-02	-11	-05	-08	-11	04	-23	12	5.34	4.32

* Decimal points omitted, $r_{.05} = .16$, $r_{.01} = .21$

11, correlated .51 with the Creativity Rating and -.50 with the Creativity Key based on the first study which was variable #12, i.e., variable 12 was the total score derived from Form A at the first administration. It is the sum of variables 71, 72, 73, and 74. The validity coefficient of .48 for this Form A total score against the Creativity Rating Criterion was considered to be very satisfactory for a number of reasons. The key was built on scientists at a different NASA research installation with different research specialities and in an entirely different part of the country. In addition, the Creativity Rating was the criterion used in the second study, yet the biographical scores based on the first study were developed and validated on the basis of the Creativity Check List. Furthermore, it will be remembered that the Creativity Check List and the Creativity Rating (the latter was administered 3 months later) correlated .69 on the one sample for the first study upon which complete data were available. Thus, in view of the unreliability which was present in the criterion data, obtaining a cross-validity coefficient of .48 at a new research center was a gratifying research finding. This, of course, implies that the Biographical Inventory's keys can be applied in various NASA research centers and that keys will not have to be built separately for each research installation. This same key based on the first administration also predicted most of the other criteria at the new research center with a satisfactory degree of validity. That is, the Form A key scores correlated .40 with the Number of Publications, .28 with the Number of Patents, .37 with G. S. Level, .38 with Knowledge of Work, .33 with Initiative, and .32 with Judgment. Although the correlations with the other criteria, Industry, Reliability, and Cooperation, were unsatisfactory, as will be discussed later, this was probably due to the inherent unreliability in these latter criterion measures.

Variables 13 through 67 are all empirical keys based upon Sample II and thus these are all cross-validity coefficients. The first four of these, variables 13 through 16, in Table 3 correspond to the sections of the Biographical Inventory as previously discussed, that is, the Developmental History Section, the Parents and Family Life Section, the Academic Background Section, and the Adult Life and Interests Section. In terms of how well they predicted the criteria, these sections are listed in terms of reverse rank order. This is somewhat in contrast to the results obtained on Form A, where Section One, the Developmental History Section, was better than Section Two, the Parents and Family Life Section. This is probably due to the fact that in Form B Section Two was shrunk considerably so that the more valid items tended to remain while Section One was expanded with new and less stable items. The small difference could, however, be due to chance. The Total Score, variable 17, which was the sum of the previous four, had a correlation of .49 with the Creativity Criterion. This is in contrast to the cross-validity coefficients which were obtained in the first study, since there, the cross validities ranged in the mid fifties for the Creativity Key. On the other hand, it will be noticed that this Total Score correlates only .01 higher with the Creativity Criterion than the Form A Key that was built on the scientists at the first research center which, as previously mentioned, had a validity coefficient of .48 against the Creativity Criterion. This indicates a surprisingly high degree of stability in the scoring keys that were developed at the first administration since it was almost as satisfactory as the key actually constructed on the basis of the responses of the scientists at the second research center. These results certainly indicate a substantial degree of stability in the scores from biographical keys as the cross-validity coefficients obtained are remarkably consistent.

The cross validity of the Publications Total Score, variable 22, against the Publications Criterion was .55. It will be noticed that the Section Four Key actually had a correlation of .56 against the Publication Criterion, slightly higher than the total score. This occasionally happens, due to the fact that other sections of the inventory are not as predictive (i.e., not as discriminatory, some may have either zero or negative validity) and the simple additive sum is not the optimum way of combining the subscores. This issue will be finally resolved and capitalized upon when factor scores are obtained from a factor analysis of the biographical items in the inventory. The same phenomenon is evident in the correlations of Patents Key Total Score against the Patents Criterion which was .22 and the Section Four score which had a correlation of .30 against the Patents Criterion. Without taking time to discuss the rest of the cross-validity coefficients in detail, a brief summary follows: for the G. S. Level Key total score predicting the G. S. Level Criterion .56; Knowledge of Work Key total score against that criterion .35; the Initiative Key total score predicting the Initiative Criterion .28; the Judgment Key total score against the Judgment Criterion .37; the Industry Key total score predicting the Industry Criterion .11; the Reliability Key predicting the Reliability Criterion .07; the Cooperation Key predicting the Cooperation Criterion -.01 and the Off-line Key predicting the Off-line Criterion .50. The comparatively lower cross-validity coefficients of the latter criteria which were collected at the second research center (Industry, Reliability, and Cooperation) are probably due to the manner in which these criteria were measured and the fact that the biographical items were primarily constructed and selected to predict other types of criteria. All of these traits were on a single form and an inspection of the ratings after they had been filled out revealed that there was comparatively little consistency in these ratings. Since the Biographical Inventory does predict the other criteria with comparatively higher cross-validity coefficients it appears, then, that these lower validities for the latter criteria are primarily a function of the criteria themselves. A more impressive argument, and an intriguing finding also, about the keys which were constructed to predict those latter criteria, Industry, Reliability, and Cooperation, is that although they did not have any appreciable cross validities against the criteria they were constructed to predict, they did have at least moderate cross validities against the more reliable criteria. For example, in Table 3 the Reliability Biographical Key, constructed on the basis of Sample II and applied to Sample I has a cross-validity coefficient against the Reliability Criterion of only .07. Yet the scores from this key correlated with the G. S. Level Criterion .53. It also correlated .48 with the Number of Publications, and .38 with the Creativity Rating. The same tendency is also evident in Sample I which adds up to an impressive argument that the Reliability Criterion Rating is unreliable. Stated in other words it appears that the key built upon this criterion has more length and thus more reliability than the criterion itself. The same tendency is also true for the keys built to predict the Industry and the Cooperation criteria.

The high cross-validity coefficients obtained from the G. S. Level Total Score Key against the G. S. Level Criterion, .64, the Creativity Criterion, .51, and the Publications Criterion, .57, argue that even higher cross validities might be obtained if more reliable criteria were available. That is, the scores from the key constructed to predict the G. S. Level Criterion predicted the Creativity Criterion and the

Publications Criterion higher than the keys which were specifically constructed to predict these criteria--these results are cross validities, not initial validities.

Variable 68, one of the correction scores and the first of the a priori predictors, correlated .21 with the Creativity Rating and -.07 with the Creativity Key. As mentioned previously these scores will be presented in detail in a later report but, for the moment, it will be noticed that this pattern of correlations would result in additional variance accounted for when added to the Form A Biographical Total Score Key. Variables 69 and 70, Rank in College and Rank in High School, respectively, generally had very low correlations with the criteria, and they correlated slightly higher with the Creativity Key Scores from the first study. These results do not argue that academic achievement is a very good indicator or predictor of success in science. These variables of course, must be interpreted with caution since they are self-reports and may not be accurate descriptions of the scientists' actual rank in college and high school. Variables 71 through 74, inclusive, were the a priori keys derived on the basis of the analysis of the scientists at the first study, as mentioned previously. Since they were designed and constructed to predict creativity it is no surprise that they do tend to predict creativity higher than the other criteria with Professional Self-confidence, Independence, General Intellectuality and the Miscellaneous Items Key being in rank order in terms of their predictive powers. Independent Intellectuality, variable 75, was also explained previously. It is sufficient to note that it had insignificant correlations with all of the criteria except for the Off-line Measure with which it correlated -.23, evidently indicating that this score might make a small contribution to the identification of those scientists that would not be accurately identified by their regular biographical scores.

Analysis of Sample II in the Form B Study. The criterion intercorrelations, validities, and standard deviations for Sample II are presented in Table 4. The variables in this table tend to parallel those discussed in Table 3, except that certain variables were eliminated from consideration in order to make more room for other predictor scores. Thus, it will be noticed that certain section scores from the various empirical keys that did not have any significant validity coefficients across the criteria were eliminated. These results were obtained through pilot studies, and hence for the more thorough analyses presented here they could be eliminated to make room for other predictor variables. The a priori keys included in the analysis of Sample II are described below.

Variable 63, Extent of Participation, was based on the rationale that extensive participation in a variety of science-like activities prior to the age of 18, such as nature study, radio, electronics, photographic processing, etc., might be related to later achievement in science. Accordingly, 17 items were selected from the Biographical Inventory which measured the frequency and extent of participation in these activities and were scored so that high participation would earn a high score.

Academic Science Performance, variable 64, was a self-report on the level of achievement obtained by each subject during high school and college in five subject matter areas, namely, biological sciences, physical sciences, mathematics, engineering, and social sciences.

Variable 65, Minimum Satisfaction, and 66, Dedication to Work, were both a priori motivation scores. The former measured the degree of achievement necessary to obtain minimum personal satisfaction in three facets of science, namely, number of publications, level of original work and level of theoretical contributions. The latter a priori motivation measure, Dedication to Work, was also made up of three items: the extent to which hard work was felt to be the basic factor of success, the degree of absorption in work, and the extent to which work activities influenced, or interfered, with other aspects of daily living.

Variable 67, Childhood Unhappiness, was of interest from a theoretical point of view, since the scientist is often characterized as one who has been reared in a rather cold emotional environment and/or his perception of his childhood is often described as being somewhat less than happy. Accordingly, this score consisted of seven items, all of which were scored in a continuous fashion to measure such characteristics of family life as the frequency of disagreements with the parents, the marital happiness of the parents, the frequency of emotional upsets during childhood, the satisfaction derived from the family life, etc.

Variables 68, 69, and 70, were the a priori keys derived from the first study of the Biographical Inventory: General Intellectuality, Independence, and Professional Self-confidence, as discussed previously.

Protestant Ethic Orientation, variable 71, is largely self-explanatory in that items were selected from the Inventory which were thought to be related to this conceptualized pattern of behavior. This score was made up of 9 items measuring such characteristics as achievement orientation of childhood friends, self-description of childhood ambition, the degree to which the parents encouraged thriftiness and saving, the importance of personal and economic success, etc.

Variable 72, The Total Times Scored, was simply the algebraic sum of the number of times each scientist was scored either positive or

Table 4

CRITERION INTERCORRELATIONS, VALIDITIES, MEANS, AND STANDARD DEVIATIONS

FOR FORM B SAMPLE II (N = 153 N.A.S.A. SCIENTISTS)*

				1	2	3	4	5	6	7	8	9	10	11	12	Mean	S. D.
Criteria	1.	Creativity Rating		—												41.01	12.40
	2.	Publications		38	—											5.14	5.84
	3.	Patents		25	07	—										.38	1.08
	4.	G. S. Level		61	07	18	—									13.02	1.90
	5.	Knowledge of Work		69	33	21	46	—								84.07	11.94
	6.	Initiative		63	09	24	31	75	—							84.09	11.73
	7.	Judgement		58	18	10	38	78	75	—						82.45	11.02
	8.	Industry		46	11	18	38	80	82	81	—					86.83	11.34
	9.	Reliability		46	27	12	33	67	65	71	81	—				88.11	11.46
	10.	Cooperation		27	11	05	25	50	34	47	42	53	—			71.58	8.58
	11.	Off-line Measure		46	-11	05	07	29	35	35	29	38	18	—		80.82	12.09
	12.	Form A Key Total Score		46	47	19	50	38	35	61	17	18	07 01	-08	—	61.10	13.83
Sample I Empirical Keys	13.	Section I	Creativity Key	-07	-02	-02	-10	-01	-03	02	-05	-02	01	11	-17	45.97	8.08
	14.	Section II	"	12	11	11	30	11	09	01	14	05	08	-17	29	52.07	2.07
	15.	Section III	"	18	18	02	13	25	24	25	13	22	11	-02	01	44.38	4.13
	16.	Section IV	"	46	50	18	53	38	34	32	10	16	08	-40	05	48.34	10.07
	17.	Total Score	"	43	48	16	48	41	27	34	21	21	12	-44	06	46.22	13.38
	18.	Section I	Publications Key	14	25	10	34	09	04	05	-01	01	08	-24	17	50.22	3.08
	19.	Section II	"	12	08	09	10	13	18	03	04	09	04	-13	28	48.45	2.34
	20.	Section III	"	14	34	-03	35	17	07	18	08	08	07	-18	32	50.49	3.13
	21.	Section IV	"	52	60	23	58	43	35	30	22	19	09	-33	04	50.38	8.47
	22.	Total Score	"	45	55	20	53	37	28	34	17	17	14	-01	74	50.86	12.44
	23.	Section I	Patents Key	-01	21	04	08	02	01	-01	-05	07	10	-03	00	51.38	3.27
	24.	Section II	"	-05	05	-17	-05	-01	-08	-01	-08	-11	-03	-18	16	52.72	1.62
	25.	Section III	"	13	-03	-07	08	05	08	06	-08	-09	10	03	10	50.58	1.94
	26.	Section IV	"	33	32	18	28	33	35	18	09	04	-05	-40	72	51.30	5.60
	27.	Total Score	"	34	31	09	23	25	19	13	01	01	02	-44	68	55.32	6.72
	28.	Section I	G. S. Level Key	14	34	-05	34	15	06	15	-40	11	08	08	04	50.84	3.02
	29.	Section III	"	14	22	-00	17	27	19	18	04	22	08	-16	21	50.56	2.47
	30.	Section IV	"	04	06	22	61	38	35	32	22	20	13	-01	62	50.70	9.58
	31.	Total Score	"	51	57	15	64	41	34	32	18	24	16	-27	75	53.28	11.94
	32.	Section I	Know. of Work Key	02	-08	06	-03	04	-05	04	-05	-04	-02	08	-00	47.70	3.18
	33.	Section III	"	16	15	-01	15	30	19	27	15	20	17	-27	44	48.70	3.38
	34.	Section IV	"	41	51	14	42	37	34	28	14	13	13	-40	72	44.21	5.90
	35.	Total Score	"	37	42	11	34	28	25	30	18	15	18	-35	71	41.02	6.82
	36.	Section I	Initiative Key	-07	06	08	05	01	-06	00	-12	-00	-02	-07	-01	48.61	3.08
	37.	Section II	"	16	08	10	05	21	13	09	30	10	11	01	14	51.80	2.06
	38.	Section III	"	12	17	02	16	17	14	22	15	25	25	-25	27	50.07	3.70
	39.	Section IV	"	44	48	16	48	40	31	34	30	18	11	-38	78	47.40	7.18
	40.	Total Score	"	41	47	19	47	42	31	35	22	24	18	-38	77	46.18	8.03
	41.	Section I	Judgement Key	-07	-04	-03	-01	04	01	-01	-11	-15	-14	-08	01	47.80	1.87
	42.	Section II	"	10	10	06	10	08	07	03	08	02	07	-08	17	51.94	2.82
	43.	Section III	"	07	08	01	00	18	30	18	14	20	13	-25	32	48.80	3.70
	44.	Section IV	"	38	44	12	40	38	31	32	18	18	10	-33	71	44.63	6.18
	45.	Total Score	"	33	38	10	32	38	32	30	18	18	10	-27	69	43.31	9.30
	46.	Section I	Industry Key	03	07	00	07	-04	-04	-02	03	-04	04	-02	04	50.74	3.03
	47.	Section III	"	02	12	-05	05	10	04	05	13	18	18	-31	28	51.13	3.72
	48.	Section IV	"	25	38	04	38	12	18	13	13	09	17	04	19	48.27	3.08
	49.	Total Score	"	18	22	03	38	18	05	09	13	10	21	-10	28	50.78	5.78
	50.	Section I	Reliability Key	-09	03	-03	11	-13	-02	03	06	04	11	01	-09	51.10	2.78
	51.	Section II	"	27	10	07	33	11	13	05	18	08	18	14	11	52.51	2.80
	52.	Section III	"	-04	10	-07	01	-06	-01	-02	00	12	06	-14	12	51.03	3.14
	53.	Section IV	"	18	28	-03	29	-02	-03	06	-01	06	11	08	08	44.38	2.83
	54.	Total Score	"	18	24	-02	34	01	00	07	07	13	10	06	10	51.09	5.80
	55.	Section I	Cooperation Key	-05	-04	02	00	02	-05	-01	-07	-08	06	08	-13	48.52	2.18
	56.	Section II	"	-10	-11	15	-12	-14	-06	-13	-06	-11	-10	10	-10	49.01	1.38
	57.	Section IV	"	02	-01	02	21	-06	-01	07	-06	-03	16	16	-14	50.28	2.78
	58.	Total Score	"	00	01	07	01	-03	-03	03	-09	-09	10	13	-13	49.06	4.83
	59.	Section I	Off-line Key	-04	-15	-08	-05	-01	-01	01	-01	01	-03	61	-41	53.17	4.38
	60.	Section III	"	-20	-18	04	-19	-21	-16	-12	-14	-13	-08	31	-51	50.63	3.47
	61.	Section IV	"	-40	-39	-19	-37	-30	-31	-21	-21	-09	-05	60	-64	47.77	10.6
	62.	Total Score	"	-30	-33	-13	-30	-24	-23	-16	-18	-07	-08	54	-63	54.78	15.4
Other Predictors	63.	Extent of Participation		-05	00	06	02	00	-04	-06	-20	-13	02	-40	35	43.18	9.7
	64.	Academic Science Performance		-05	13	02	10	03	04	03	12	15	14	-38	34	13.53	2.6
	65.	Minimum Satisfaction		25	32	17	17	16	16	02	10	-02	06	-32	56	12.09	3.5
	66.	Dedication to Work		27	09	03	21	23	19	15	03	00	-13	-07	32	10.60	1.9
	67.	Childhood Unhappiness		14	11	09	05	04	09	-01	-01	03	-04	-10	22	16.34	3.
	68.	General Intellectuality		19	32	15	30	18	19	15	06	08	01	-68	78	51.26	4.7
	69.	Independence		40	30	15	33	31	31	34	19	10	02	-30	68	48.39	3.
	70.	Professional Self-confidence		50	42	19	55	42	38	34	20	20	12	-42	90	48.18	6.
	71.	Protestant Ethic Orientation		-07	18	02	06	-01	-02	01	09	06	06	-34	27	27.11	4.
	72.	Total Times Scored		03	04	-09	-10	00	07	05	-01	03	-03	-10	13	45.05	5.
	73.	Age		35	44	15	59	23	13	19	16	23	22	11	27	38.05	8.
	74.	Father's Education Level		09	05	00	-05	08	-09	-10	-10	-07	-07	05	03	46.84	21.
	75.	Subject's Education Level		38	29	-07	30	32	12	12	17	08	-02	-02	41	22.71	11.

* Decimal points omitted, $r_{.05}$ = .16, $r_{.01}$ = .21

negative on the Sample I, Form B Creativity Key. This variable was the same as in the Form A study of the Biographical Inventory. Variables 73, 74, and 75 are largely self-explanatory, measuring the age of the subjects, and the amount of schooling completed by the father and by the subject.

The results from the analysis of Sample II are presented in Table 4. Generally speaking the intercorrelations and validity coefficients for Sample II tend to follow closely the results already discussed for Sample I. Since much of the information in this table is similar to the results of the table previously discussed, they will be treated in a summary fashion pointing out only those of special interest.

The Total Score on the Biographical Inventory from the first study, variable 12, correlated with the Creativity Criterion Rating .46. It will be remembered in the previous sample of the Form B study that the correlation was .48, thus, the average cross validity on the two samples was .47. This is additional evidence of the very stable nature of the Biographical key scores. The cross-validity coefficients of each of the empirically keyed scores against their appropriate criterion was as follows: the Creativity Key total score against the Creativity Rating Criterion .43, the Publications Key total score against the Publications Criterion .55, the Patents Key total score against the Patents Criterion .09, the G. S. Level Key total score against the G. S. Level Criterion .64, Knowledge of Work Key total score against the Knowledge of Work Criterion .38, the Initiative Key against the Initiative Criterion .31, Judgment Key total score against the Judgment Criterion .30, the Industry Key total score against the Industry Criterion .06, the Reliability Key total score against the Reliability Criterion .13, the Cooperation Key total score against the Cooperation Criterion .06, and the Off-line Key score against the Off-line Criterion .54. Comparing these cross-validity coefficients to those of the Form A Total Score, variable 11, it will be noticed that the Form A Total Score from the first study did almost as well in predicting the criteria on the sample of scientists in the second study as did the keyed scores which were specifically constructed to predict the criteria at the second research center. In fact scores from the Form A Key correlated higher with the Creativity Rating Criterion (.46) than did scores from the Form B Creativity Key (.43). This is again evidence that the keys built on the samples of scientists in the first study do have a high degree of stability and can be used at more than one NASA research center.

Variable 63, the Extent of Participation Score, did not work as anticipated, since it correlated .05 with the Creativity Rating Criterion and also had essentially zero correlations with the rest of the criteria. There are a number of ways in which these results can be interpreted. To the extent that this score accurately reflects the extent of participation in a variety of science-like activities, then the results do not speak very well for Science Fairs and other activities which encourage an early participation in science. That is, the present results do not indicate that such participation is of a marked benefit. Of course neither do they indicate that such participation is in any way detrimental to the development of later competence in science. In any case it is an interesting question, one which should be followed up with additional research on Science Fairs and the people who participate in them. This Extent of Participation Score did correlate moderately high with the Total Biographical Score, .35, and with the Off-line Measure, -.40. This could indicate that the group of items selected for this measure were subject to a response set on the part of some of the scientists filling out the Biographical Inventory. That is, there was evidently a tendency for those who scored high on the Biographical Inventory to perceive these items either consciously or unconsciously as being indicative of the characteristics which should describe good scientists. Speaking in more traditional psychometric language, those scientists who scored high on this score would tend to be identified as false positives since their biographical score was higher than their criterion score. In terms of the correction score which will be discussed in a later report they could be thought of as exaggerators. Another way of viewing this score and the results that were obtained is that the scientist filling out the Biographical Inventory perceived these items the same way as the investigators did in designing this score. That is, they evidently felt these items would be characteristic of how successful scientists would respond or should respond to the Biographical Inventory.

Academic Science Performance, variable 64, which was a self-report on the level of achievement by each subject during high school and college in five areas, biological science, physical science, mathematics, engineering, and social sciences, had essentially zero relationships with all of the criteria. But once again a moderately high correlation was obtained between this variable and the Biographical Score as well as the Off-line Criterion measure. To the degree that these self-reports are accurate they indicate, then, that achievement in high school and college in those areas had essentially a zero relationship to the criteria of success in science used in this study, but that obtaining A's or high grades in these areas was related to the score that was obtained on the Biographical Inventory. Perhaps this is a facet of how they perceive themselves, that is, the scientists who were more successful during their academic training also perceived themselves as being more successful as they filled out the Biographical Inventory. In other words they have a high self-concept, higher than what the criterion measures would indicate as realistic. These findings again raise questions and doubts about the appropriateness of our educational system. There have been enough studies which have conclusively demonstrated that, at best, the relationship between grade-point average and performance on the job is low, if statistically significant. These are not new findings, then, but they do add substance to the claim that our educational system needs to be improved, if grades are to be really representative of the type of performances that are needed in the world of work. (Taylor, 1958, 1963; Business Week, 1962; D. Taylor, 1958).

Variable 65, Minimum Satisfaction, and variable 66, Dedication to Work, were both short *a priori* motivation measures. The Minimum Satisfaction Scale had rather consistent correlations across the criteria, correlating .25 with the Creativity Rating and .32 with the Number of Publications. The Dedication to Work Scale, while not having validities quite as high as the Minimum Satisfaction Score, was less related to the total score of the Biographical Inventory, correlating only .32 in contrast to the .56 correlation between the Minimum Satisfaction Scale and the Biographical Total Score. It is of interest to note that this Dedication to Work score correlated highest with Creativity, next highest with Dedication to Work, and then the G. S. Level Criterion. It had essentially zero relationships with the Number of Publications and Patents. This suggests that this score measures a task orientation and actual involvement in work which does not find expression in written reports. It seems to involve work for its own sake.

Varialbe 67, Childhood Unhappiness, did not have any significant correlations with the criteria although the correlation with the Creativity Rating was close to significant. It did correlate significantly with the Biographical Score, .22. Perhaps if this scale were lengthened the validity might increase to the point where it could be demonstrated that these items do at least have a theoretical significance if not a very practical significance for identification of scientific talent and other criteria of success in science.

Variables 68, 69, and 70, the three keys based on the first study of the Biographical Inventory, tended to have their usual pattern of significant validity coefficients with Professional Self-confidence being the most valid, followed by Independence and then General Intellectuality. The Professional Self-confidence Key correlated higher with the Creativity Criterion (.50) than did the Form B Creativity Key Total Score (.43). In terms of variance accounted for by each of these keys, there is a 7% difference, with the key from a different research center being superior to the key that was developed and cross validated to maximize the prediction of the Creativity Criterion at the second research center. It is of interest to note that the General Intellectuality Score correlates higher with the Number of Publications than it does with any other criterion. It also correlates higher with the Off-line Criterion Measure, -.59, than do the other two a priori keys.

Variable 71 was the Protestant Ethic Orientation, an a priori key. With the exception of a significant correlation with the Number of Publications, scores from this key did not have any other significant validities, except with the Biographical Key scores and the Off-line Measure. These results might be interpreted in two ways. First, the key may actually measure a Protestant Ethic Orientation since the significant correlations with the Number of Publications would be positive evidence that one has achieved, or better yet, demonstrated that he is among the divinely chosen. If this rationale is realistic then the low negative correlation with Creativity is of interest since this would be consistent with the rationale that it is the demonstrable products or achievements that are important, and that an above-average level of creativity is not as important as positive demonstrable products of one's effort; in fact, for some people such an orientation might actually detract from an orientation toward creativity. Second, this key could also be interpreted as being similar to the Extent of Participation Score discussed above. That is, it could be measuring items that are primarily associated with a response set and as such this would explain the Correlation of .27 with the Creativity Key and the -.07 correlation with the Creativity Rating.

Variable 72, the Total Times Scored on the Sample I Form B Creativity Key, did not produce the same results as an analogous score derived in the Form A Study; in fact this type of score in the Form A study correlated positively with most of the criteria, whereas in the Form B study it had essentially zero correlations, not only with all of the criteria, but also with the Form A Total Score (Variable 12) and with the Off-line Measure. No satisfactory explanation has yet been suggested for these results--apparently the results are inconsistent for this score as its keys is developed on each new sample. It may reflect to a small extent the "research climate" at various NASA research centers. This topic will be treated more thoroughly in a discussion of variable 75.

Variable 73, Age, was included to determine its relationship to the criteria and to the biographical scores. Generally speaking, age correlated moderately high with most of the criteria but only .22 with the Form A Key Score. Thus, scores on the Form A biographical key are not largely predicting age even though age is related to the criterion measures.

Variable 74, Father's Education Level, did not have any significant correlations with any of the other variables, thus indicating that for the scientists in this sample, the education of the father had no influence on the achievement of the son. This does not tell us anything, of course, about any effects of the education of the father on the son's attending or graduating from college or choosing science as a profession.

Variable 75, Subject's Educational Level, had a pattern of surprisingly high correlations across the criteria and with the Form A biographical score. These results are in contrast to those of Taylor et al. (1961). In this earlier study, they found that years of education had either low or zero correlations with the multiple criteria of success in science studied of which there were 14 essentially independent dimensions. In view of these apparently contradictory results, additional evidence was gathered together to try to clarify this issue. Additional information was available from the item analysis results of the Biographical Inventory. In the Form A study, the level of education also had some significant correlations with the criteria but generally they were not as high as in the Form B study. In the Form A study, obtaining a B.A. or B.S. degree or less had a low but negative relationship to the criteria while completing some graduate work but no graduate degree, or completing an M.A., or M.S. degree, or completing some graduate work beyond the master's degree all had essentially zero relationships with the criteria. Completing the Ph.D. degree, however, was positively related to the criteria, even though only nine per cent of the sample in the Form A study had completed the Ph.D. degree.

Another source of information on this topic comes from a recently completed study by Taylor, Cooley, and Nielsen (1963). In this study, which will be more thoroughly discussed in a later section of this report, high school and college students who participated in NSF-supported summer science training programs were given a revised form of the Biographical Inventory and other selected predictor instruments in an attempt to "carry back" and validate instruments developed on mature scientists to those younger age groups. In this high school study both age and grade level tended to correlate either zero or negative with the 15 criteria of achievement in science that were used. Yet with age and grade level essentially "partialled out," the Biographical Inventory with keys built on NASA scientists proved to be one of the best predictors used in the study. Thus, it appears that although the Form A Biographical Inventory Key is moderately related to level of education on this sample, this relationship appears to be at least partially contingent upon the sample studied. And it also appears evident that the key contains much additional variance which is related to criteria of science achievement other than that which would be identified by level of education. It is also interesting to note that the keys constructed on the scientists in the Form B study have in recent research turned out to be "academic" keys. That is, in a study by Bunderson, Rigby, and Taylor, (1963), the Form B keys turned out to be the best predictor of combined academic and research performance in graduate school. Also, in the NSF study mentioned previously the Form A keys worked better in the research participation-type programs, while the Form B keys worked better in the classroom type of program. These findings suggest that there may be interesting "climate" differences between the two research centers studied.

Chapter 4

THE FORM C STUDY OF THE BIOGRAPHICAL INVENTORY

Another 300 item form of the Inventory was constructed in which the best items from the previous studies were used and some new items were added. This form (Form C) was administered to 769 scientists at a third NASA research center where ninety-seven per cent of the Biographical Inventories distributed were either completed or otherwise accounted for, still another excellent example of the cooperation which has been so characteristic of this study. The scientists at this third center conduct research in aerodynamics of re-entry vehicles, structures and materials for space vehicles, aircraft aerodynamics, fundamental plasma physics, and a wide variety of other areas.

In contrast to the other two research centers visited, there was no existing rating procedure at this center for the evaluation of scientific personnel. Promotions were handled by means of letters of recommendation and by meetings of those concerned. Thus, the criterion measures collected at this center may have been influenced by this comparative lack of rating experience. In addition to collecting information on the Number of Patents, the Number of Publications, and the GS Level attained, other criterion measures were collected from supervisors. These instruments were constructed by the investigators to evaluate the following traits: Quantity of Work, Skill in Getting Along with People, Creativity, and an Overall Evaluation. These criterion rating forms resembled those constructed for research purposes at the two previous centers. A sample of these forms is presented in the Appendix.

The procedures followed at this third NASA center were again the same as in the previous studies. Criteria were collected from supervisors, the biographical inventory was administered to all the scientists, and after the total sample was divided into two sub-samples of 390 and 379, an item alternative analysis was carried out in a double cross validation design.

Since all of the working items from the previous forms were retained in Form C, the scoring keys from the previous studies were also included in the analysis, that is: the General Intellectuality Key, the Independence Key, the Professional Self-confidence Key, and the Form A Total Score from the Form A Study, together with the Form B Total Score Key. The latter included all of the items to which at least 5 per cent of the total Form B sample responded and which also had an average correlation of plus or minus twenty or greater across the two sub-samples in the Form B Study with the Creativity Criterion.

Results for Main Samples. Criterion means, standard deviations, intercorrelations, and cross validity coefficients for the keys from the previous studies together with the various new empirical keys constructed at this research center are shown in Table 5 for Sample One. Table 6 shows the corresponding results for Sample Two. The Criterion intercorrelations are similar to those in the previous studies, with two exceptions. In the Form B Study where Number of Publications was also retained as a separate criterion, its correlation with the Creativity Rating in one sample was .46 and in the other sample .38. However, in both of the samples in the Form C Study the corresponding correlations were considerably lower, .17 in Sample One and .21 in Sample Two with the Creativity Criterion. In addition, the G. S. Level Criterion correlated considerably lower with the Creativity Criterion in the present study than previously. For example, in the Form B Study, G. S. Level correlated .61 in one sample and .55 in the other sample with the Creativity Criterion, while in the Form C Study the correlations of G. S. level with Creativity were .32 in Sample One and .36 in Sample Two.

These results evidently suggest that the supervisors in their ratings did not take into consideration, to the extent that was done in the previous

Table 5

CRITERION MEANS, STANDARD DEVIATIONS, INTERCORRELATIONS AND CROSS VALIDITIES FOR SAMPLE ONE IN THE FORM C STUDY OF THE BIOGRAPHICAL INVENTORY*

(N = 390)

No.		1	2	3	4	5	6	7	Means	S.
	SAMPLE ONE CRITERIA									
1	Quantity of Work	--							9.21	2.5
2	Skill with People	65	--						8.92	2.3
3	Creativity of Work	69	58	--					8.35	2.6
4	Overall Work Performance	85	72	83	--				9.24	2.4
5	Patents	08	01	13	08	--			.23	.8
6	Publications	19	21	17	20	11	--		6.60	7.3
7	G. S. Level	27	31	32	33	13	59	--	11.91	2.2
	SAMPLE TWO KEYS (Form C Study)									
8	Quantity of Work Total Score	33	31	41	38	16	28	31		
9	Skill with People Total Score	32	36	32	34	05	29	38		
10	Creativity of Work Total Score	27	24	41	34	18	21	23		
11	Overall Work Perf. Total Score	31	29	41	37	17	27	28		
12	Patents Total Score	20	17	31	26	17	28	33		
13	Publications Total Score	24	24	33	29	22	59	66		
14	G. S. Level Total Score	21	24	28	24	19	57	69		
	OTHER PREDICTORS									
15	General Intellectuality	13	11	17	17	17	13	14		
16	Independence	22	15	28	22	16	13	16		
17	Professional Self-confidence	28	27	41	35	17	30	29		
18	Miscellaneous	13	09	10	14	-01	23	24		
19	Form A Total Score	28	24	38	33	21	26	27		
20	Form B Total Score	14	15	26	18	11	18	20		

* Decimal points omitted. $r_{.05} = .10$; $r_{.01} = .13$.

Table 6

CRITERION MEANS, STANDARD DEVIATIONS, INTERCORRELATIONS AND CROSS VALIDITIES FOR SAMPLE TWO IN THE FORM C STUDY OF THE BIOGRAPHICAL INVENTORY*

(N = 379)

o.		1	2	3	4	5	6	7	Means	S. D.
	SAMPLE TWO CRITERIA									
1	Quantity of Work	--							9.01	2.64
2	Skill with People	62	--						9.03	2.47
3	Creativity of Work	64	51	--					8.43	2.45
4	Overall Work Performance	78	69	80	--				9.20	2.33
5	Patents	05	02	14	10	--			0.38	1.20
6	Publications	15	14	21	21	22	--		6.62	7.42
7	G. S. Level	25	30	36	34	19	62	--	11.81	2.26
	SAMPLE ONE KEYS (Form C Study)									
8	Quantity of Work Total Score	31	21	40	37	26	35	38		
9	Skill with People Total Score	30	24	36	35	21	40	45		
0	Creativity of Work Total Score	27	17	41	36	23	31	34		
11	Overall Work Performance	30	20	40	38	24	35	38		
12	Patents Total Score	09	-02	27	18	24	32	29		
13	Publications Total Score	17	20	27	27	23	66	69		
14	G. S. Level Total Score	16	21	29	25	19	59	72		
	OTHER PREDICTORS									
15	General Intellectuality	10	07	24	19	18	20	18		
16	Independence	10	04	27	16	20	17	13		
17	Professional Self-confidence	31	18	40	38	19	31	31		
18	Miscellaneous	11	16	14	19	10	28	28		
19	Form A Total Score	23	13	39	32	23	29	27		
20	Form B Total Score	16	12	28	23	13	16	18		

* Decimal points omitted. $r_{.05} = .10$; $r_{.01} = .13$.

2

studies, what the scientist had accomplished in terms of publications and in terms of the level of achievement that he had obtained within the organization. Or alternately stated, the ratings in the present study had less halo and were more independent of other kinds of performance, thus they may reflect a somewhat different conception concerning creativity than evidently was the case in the previous studies. Also, it may be that people were promoted at the other centers more in terms of their creativity than at this one.

The cross validity coefficients of the various new empirical keys in Sample One were as follows: The Quantity of Work Key against the Quantity of Work Criterion .33; the Skill with People Key against the Skill with People Criterion .36; the Creativity Key against the Creativity Criterion .41; the Overall Work Performance Key against the Overall Work Performance Criterion .37; the Patents Key against the Patents Criterion .17; the Publications Key against the Publications Criterion .59; and the G. S. Level Key against the G. S. Level Criterion .69.

The keys derived from the previous studies, while generally producing significant correlations, were nevertheless somewhat lower than what had been obtained previously. For example, in Sample One the Form A General Intellectuality Key correlated only .17 with Creativity; the Independence Key with the Creativity Criterion, .28; and the Professional Self-confidence Key with the Creativity Criterion, .41.

For Sample Two, the cross validity coefficients were very similar. The new Form C Quantity of Work Key cross validated against the Quantity of Work Criterion, .31; the Skill with People Key against the Skill with People Criterion, .24; the Creativity Key against the Creativity Criterion, .41; the Overall Work Performance Key against the Overall Work Performance Criterion, .38; the Patents Key against the Patents Criterion, .24; the

Publications Key against the Publications Criterion, .66; and the G. S. Level Key against the G. S. Level Criterion, .72.

The magnitude of the cross validities on both samples against the last two criteria of Publications and G. S. Level is extremely high and certainly worthy of special notice.

The results obtained from applying the keys from the previous studies to the scientists in Sample Two of the Form C Study were similar to those obtained with Sample One, Professional Self-confidence again being the most valid scoring key. It is significant to note, however, that the Professional Self-confidence Key yielded a level of prediction which was very comparable to the Form C Creativity Key which was constructed specifically to predict the Creativity Criterion at this research installation. There has been a consistent trend across these studies for the keys constructed in the Form A Study (especially for the Professional Self-confidence Key, but also for the Form A Total Score as well) to correlate with the criteria at a different NASA research center as highly as the keys constructed specifically to predict those criteria.

Subsample Analyses. Since this was the largest sample of scientists yet studied and since the scientists at this research center are engaged in widely ranging types of research problems, a different type of analysis involving an organizational breakdown into various sub-samples of more homogeneous performance areas was carried out for the Form C Study. The scientists were placed into subgroups according to two criteria of classification: the area of research performance and the confidence expressed by supervisors in giving their ratings. The resulting five subgroups with a brief description of each are listed in Table 7. It is apparent from inspection of groups one, two, and three, that this subgrouping has resulted in three quite separate areas of research performance. Scientists who were engaged only in fluid and flight mechanics were drawn from each of the first three groups and placed into group four, which is therefore the most narrowly defined group. Those in group five were so placed because, in contrast to those not placed in this group, their supervisors were more confident about the accuracy of their appraisals of the scientists' performance.

Each of the five subgroups was then analyzed separately according to the procedures followed in the previous studies which involve an item alternative analysis carried out in a double cross validation design on two subsamples of each group.

Table 7

Organization Breakdown
for the Form C Study

Total Sample (N=379, 390)	The Total Sample, as in the previous analyses. Cross-validity coefficients listed under this group are an average of "Sample One" and "Sample Two."
1. Group One (N=111, 102)	Scientists in this subgroup are engaged in Analysis and Computation, Instrument Research, and Theoretical Mechanics.
2. Group Two (N=136, 139)	Scientists in this subgroup are engaged in Applied Materials and Physics, Dynamic Loads, and Structures Research.
3. Group Three (N=143, 138)	Scientists in this subgroup are engaged in Aero-physics, Aero-Space Mechanics, and Full Scale Research
4. Group Four (N=175, 184)	Scientists in this subgroup are engaged only in fluid and flight mechanics.
5. Group Five (N=300, 323)	Scientists were placed in this subgroup on the basis of the greater confidence expressed by their supervisors in giving their ratings.

Results for Subsamples. The results of these analyses are presented in Table 8. Although few of the obtained relationships are significantly different from the Total Sample analysis and none is significantly different in the positive direction, there are several interesting findings. The

Table 8

Average Cross Validities for Appropriate Keys Across the Two Odd-Even Sub-samples for each Organizational Group and Each Criterion in the Form C Study of the Biographical Inventory

Organizational Groups	Sample Size	Criteria						
		Quantity of Work	Skill with People	Creativity of work	Overall Work Performance	Patents	Publications	G. S. Level
1. Total Sample	(N=379, 390)	32	30	41	38	20	62	70
2. Group One	(N=111, 102)	18	28	30	28	14	39*	60*
3. Group Two	(N=136, 139)	22	26	34	34	16	57	62*
4. Group Three	(N=143, 138)	36	21	44	39	26	52*	54*
5. Group Four	(N=175, 184)	34	34	48	42	21	59	63*
6. Group Five	(N=300, 323)	34	34	44	38	23	57	63*

*Cross-validities which differ significantly from the Total Sample results at the .05 level, as computed from Fisher's Z function.

Note: decimal points are omitted.

analyses of Groups One and Two resulted in lower cross validities than in the Total Sample for all seven criteria. Group Three was interesting because the cross validities were higher for Criteria one, three, and four (all supervisory ratings) and lower for the other four criteria, with the results for Number of Publications and G. S. Level being significantly lower from the Total Sample beyond the .05 level of confidence. The analysis of Group Four resulted in cross validities for each of the four supervisory rating criteria and the Number of Patents Criterion which were higher than those obtained in the Total Sample analysis. The cross validities for this group on the other two criteria were lower than in the Total Sample, with the difference for the G. S. Level Criterion being significant. There was also an increase in the cross validities across three of the four supervisory ratings (the other cross validity being equal to that of the Total Sample) and for the Number of Patents Criterion. As in the Group Four analysis, the cross validities for the Number of Publications and G. S. Level Criteria were lower than in the Total Sample, the latter being significant.

The rationale behind this type of analysis was that supervisory ratings should be more reliable and predictable when the performances being rated are more alike in a sample than when many different types of performances are being rated in the same sample. Inspection of Table 4 shows this to be the case with Group Four on Criteria one, two, three, and four (the supervisory ratings). Although none of these cross validities was significantly different from the Total Sample, each of them was higher than those obtained in the Total Sample and a trend here is quite apparent. This agrees with our rationale because, it will be recalled, Group Four consists of only those scientists who are engaged in fluid and flight mechanics and thus their tasks and performances are more nearly alike than the other groups.

There is also the same kind of trend in Group Five where there is an increase in three of the four cross validities (the other being equal to the Total Sample) with the supervisory ratings. This was also expected since those placed in Group Five were scientists for whom raters were more confident of their ratings and presumably these ratings were more reliable and more valid than those for the other groups.

It should also be noted that each of the group's cross validities on the G. S. Level Criterion was significantly lower (at the .05 level) than that for the Total Sample. Apparently the magnitude of this relationship depends more upon the number (perhaps the variety) of subjects than upon homogeneity of the performance tasks. The same can be said for the Number of Publications Criterion since all of the groups' cross validities were lower than that for the Total Sample, with those differences for Groups One and Three being significant.

Summary of the Form C Study. The results obtained in the Form C Study were among the best in certain areas and in other areas among the worst (though still clearly significant) that had been obtained across the three NASA centers. The cross validity coefficients of .59 and .66 in predicting Publications represent a remarkable degree of prediction as do the cross validities of .69 and .72 in predicting G. S. Level. On the other hand, the cross validity coefficients in predicting the various supervisory ratings were less satisfactory than what had been obtained previously. However, the cross validity coefficients were all significant far beyond the .01 level of confidence and these results, if used in a selection situation, could still make an important contribution to the identification of scientific talent.

Chapter 5

The Form C-1 Study of the Biographical Inventory

All of the previous studies with the BI have been concurrent validity studies, providing some indication but little actual information relative to its validity in a predictive (longitudinal) situation. Therefore, the present study was undertaken to determine how well the BI would perform in a follow-up study.

Although little has been done in the way of comparison between concurrent and predictive validities for biographical inventories, a study was done on aptitude-like tests in which the concurrent validities were always higher, most of the differences ranging from .06 to .09 (Alf, 1963). These results led to the conclusion that for this type of test, a shrinkage of at least .05 could be expected in switching from a concurrent to a predictive validity study.

Alf's study, however, was designed in a way that would lead to obtaining similar results in both the concurrent and predictive validities. The same sample was studied, being tested before starting school and then retested at the later time when the criterion of school grades became available. Even though the criterion was the simpler one of school grades, none of the validities were nearly as high as the concurrent validities which have been reported for the BI. Thus, with the identical criterion on the same sample, it would be expected that differences in validity would almost be minimized. Nonetheless, there were consistent differences always in favor of the concurrent validities.

Description and Procedure of the Form C-1 Study. Each of the NASA research centers participated in the Form C-1 study by having all of their newly hired research professionals complete Form C-1 of the BI as they reported for work, and just over a year later criterion data were obtained on the performance of the research professionals. Form C-1 is identical to the 300 item Form C, with the Form C-1 designation being used here to differentiate between the two respective studies.

Since the method of administering the BI was carried out over a time period of a few months, a psychologist from the University of Utah did not visit the NASA centers to help the personnel directors of the various centers with the administration of the inventory. For the first time, also, criterion data were collected entirely by personnel within the official system, not by visiting researchers who more obviously were collecting the data to take elsewhere for research purposes only. The extent to which this procedure may have fostered an attitude that was less cooperative than what might have been obtained if a psychologist had visited the centers to inform and explain the purposes and procedures of the research is not well known. This may have been especially important in the case of the supervisory ratings, since in the previous studies the visiting psychologist spent some time with most of the participating supervisors describing the importance of and the procedures involved in obtaining accurate personnel evaluations. In a previous study (Sessions & Taylor, 1961), evidence was presented that personnel in the official system could administer tests at least equally well as visiting researchers, but the criterion data collected by members of the official system were not as good as those collected by visiting investigators which more obviously was to be used for research purposes only.

The procedures followed in the administration of Form C-1 were identical to those used throughout all of the former studies, except for the instructions. In contrast to the previous studies in which the scientists were told that all data collected were for research purposes only, the scientists participating in this study were informed that while their specific answers to each question would be kept strictly confidential, the scoring of the BI would be related to their future work performance. Only afterwards were they informed that the data would be used for research purposes only. The intent in using these instructions was to make the administration of Form C-1 of the BI more comparable to an actual hiring situation and still maintain the cooperation of all those involved.

No performance measures such as existing ratings, number of publications, or G. S. Level, which had been obtained in the previous studies, were collected for the scientists at any of the centers engaged in the Form C-1 study. Thus, the only criteria collected were ratings on scales constructed by the investigators and completed by each scientists' supervisor after each scientist had been on the job just over a year. These scales (samples of which are presented in the Appendix) consisted of three supervisory ratings: one on Overall Work Performance, one on Creativity of Work, and a completed Creativity Check List, from which two scores were derived, namely a Creativity Score and a measure of Likeability. Also a fifth criterion measure was obtained by converting both the Creativity Rating score and the Creativity score from the Check List to standard scores at each of the centers and obtaining the arithmetical mean of the two for each scientist.

It can be noted here that the time period covered by the criterion performance measures in this study is extremely short compared to that covered by the performances measures in previous studies. This could lead to a tendency for the supervisors in this study to rate the scientists more on likeableness initially than would be true later when the scientists' actual productivity and creativity would become more apparent through their job performances and products. The extent to which this factor can influence the accuracy and reliability of performance measures is indicated in later sections of this report where the results of the several analyses are discussed.

Complete data on the Biographical Inventory, the two Supervisory Ratings, and the Check List were obtained for 622 scientists from six different research centers. A return of this magnitude is truly outstanding since a researcher from the University of Utah did not participate in the collection of the data, thus illustrating once again the excellent cooperation of the NASA personnel in providing such returns by mail.

Two kinds of analyses were carried out in the Form C-1 study, both utilizing high speed computer techniques. In the first analysis, scoring keys derived from previous studies were applied to the BI responses of each scientist. The resulting scores were then correlated with the criterion measures to obtain cross validity coefficients for each of the centers and for the total sample. The second kind of procedure involved an item alternative analysis on the total sample carried out in a double cross validation design (the same procedure used in the former studies).

Scoring Key Analysis in the Form C-1 Study. Seven previously developed keys were used in scoring the BI responses. Since an elaborate description of their content and development is presented **earlier in this** report as discussed in the Form A and Form B studies and in Ellison (1964), only a brief account will be given here. The keys used were: Creativity I (Minimum Weight), Form A Total Score, Creativity II (Maximum Weight or Professional Self-confidence), Modal Response, Off-Line, False Modesty, and Exaggeration. In general, the keys were developed by examining the item alternative analyses of the previous studies and retaining those item alternatives which resulted in a cross validity coefficient of at least .20 with the appropriate criterion and to which at least 5 per cent of the sample responded. The Creativity I or Creativity Minimum Weight Key was constructed in the above manner based on the Creativity Criterion measures in the previous studies. It differs from the Creativity II Key in that the minimum validity cut-off level for the retention of item alternatives was lowered so that more item alternatives were scored. Also the Creativity II Key was originally an a priori key which was refined through empirical analysis. The Form A Total Score Key was derived from the procedure described above in the Form A study of the Biographical Inventory.

The remaining four keys are correction scores designed to adjust distortions in the BI responses. That is, the Modal Response Key was designed to account for the BI scores of those scientists who tend to respond in the particular modal range of the various BI items, and thus their responses do not reflect an accurate BI score in relation to their criterion score. The Off-line Key was designed to account for the degree of error (regardless of direction) that occurred in trying to predict the criterion distribution from the distribution of the BI predictor scores for a sample of scientists. This Off-line Key thus was an attempt to identify those scientists who, in psychometric language, would be called false positives--those who had a BI predictor score larger than their criterion score, and false negatives--those who had a predictor score lower than their criterion score. Finally, the Exaggeration Key and the False Modesty Key were derived from the Off-line Key. Those scoring high on the Exaggeration Key consisting of scientists identified as false positives--their self descriptions being somewhat inflated as judged in light of their criterion measures. Those scoring high on the False Modesty Key were identified as false negatives--those scientists who were conservative and tended to describe themselves, their backgrounds, and experiences in an overly modest and humble fashion in relation to their actual performance.

With seven different predictor and correction score keys, various combinations can be utilized in partial and multiple correlational methods for maximum prediction of the criteria as presented in Guilford (1954). A procedure of this nature should result in greater determination of the dimensions involved in the variables under investigation.

The data analysis for this part of the Form C-1 study was completed at the Western Data Processing Center at Los Angeles, California and was carried out with high speed electronic computers. The computer program scored each subjects' BI responses according to the seven above mentioned keys, correlated each resulting score distribution with the criterion measures, intercorrelated the criterion scores and also intercorrelated the keyed scores. Needless to say, this procedure added a great deal to the speed and overall efficiency in determining the results of this study.

Results of the Combined Sample Analysis (Sample I). The procedure described above was first applied to the combined sample of 622 scientists from six different research centers. Combining the scientists for analysis in this manner was justified since computation indicated no significant differences from center to center in all of the mean criterion scores and all seven of the mean keyed scores. There is also some justification for this from previous comparable results across centers indicating that if results obtained here were again positive, this would suggest broad applicability of the scoring keys and the research procedures.

The results of the Combined Sample Analysis are presented in Table 9. The first three variables listed are criterion measures. These three criteria were used in the Total Sample Analysis because these were the three measures common to all six research centers involved. That is, certain criterion scores were obtained at several of the centers which were not obtained at others. The interrelationships between these three

Table 9

CRITERION INTERCORRELATIONS, KEY SCORE INTERCORRELATIONS, MEANS, STANDARD DEVIATIONS, AND CROSS VALIDITIES FOR THE TOTAL SAMPLE (SAMPLE I) IN THE FORM C-1 STUDY*

(N = 622 NASA SCIENTISTS)

	1	2	3	4	5	6	7	8	9	10	Means	S. D.
1. Overall Work Performance	--										9.56	2.06
2. Creativity Rating	71	--									8.61	2.28
3. Creativity Check List	69	74	--								49.49	10.26
4. Creativity 1 Key	09	15	17	--							98.79	7.46
5. Form A Total Score Key	13	18	20	88	--						108.30	10.49
6. Creativity 2 Key (Prof. Self-Conf.)	12	17	20	72	85	--					100.36	5.40
7. Modal Response Key	07	04	05	-14	10	-09	--				109.14	2.71
8. Off-Line Key	-04	-06	-10	-45	-21	-43	-01	--			100.38	3.8
9. False Modesty Key	-06	-06	-12	-29	-13	-32	-15	84	--		104.22	2.3
10. Exaggeration Key	-02	01	02	48	25	43	-14	-78	-38	--	96.08	2.3

Decimal points omitted. $r_{.05} = .08$; $r_{.01} = .11$

universally used measures are quite high, generally exhibiting more common variance than had the criterion scores in the previous studies of the biographical inventory, possibly due to the shorter time on the job and the less differences across performances yet noted by supervisors.

Variables 4, 5, and 6 are predictor keys developed from previous studies so that, as well as for the last four correction keys, the correlations between these measures and the criteria are all cross validity coefficients. The Creativity I Key correlated .09 with the Overall Performance Criterion, .15 with the Creativity Rating, and .17 with the Creativity Check List. As expected, this key was more effective in predicting the Creativity Criteria than other types of performance. The Form A Total Score Key correlated .13 with the Overall Performance Criterion, .18 with the Creativity Rating, and .20 with the Creativity Check List. Again the Creativity Criteria were predicted to a higher degree than was the Overall Performance Criterion, but there was some increase in its prediction with this Form A Key over that found for the Creativity I Key. The Professional Self-confidence Key (Creativity 2) exhibited about the same pattern, correlating .12 with the Overall Performance Criterion, .17 with the Creativity Rating, and .20 with the Creativity Check List. The key scores were quite highly interrelated, the intercorrelations ranging from .72 to .88. The validities reported here are much lower than those obtained in the predicting of similar criteria in previous studies, while the key score intercorrelations here are somewhat higher.

Although eight out of nine of these cross validities are significant beyond the .01 level of confidence, the magnitude of these relationships are considerably lower than what had been obtained in previous studies. Some decrease could be expected in a follow-up study of this nature, but somewhat better results were expected since these predictor keys had held up very well in previous studies, even from one center to another. There are several possible explanations for these results, perhaps the most valid of which concerns the criterion measures. This avenue of explanation will be referred to throughout this report when relevant evidence is discussed, and a synthesis will be attempted along with other considerations when all the evidence has been presented.

Variables 7 through 10, the correction keys, did not perform as well as was expected. These keys were constructed to perform either as suppressors or to identify any possible additional criterion variance not accounted for by the predictor keys. In order to perform this function, the correction key must either act as a suppressor or must itself contribute something to the prediction of the criterion. Inspection of Table 5 shows this to be the case with Variable 7, the Modal Response Key, which has positive correlations with all three criteria and negative relationships with variables 4 and 6, which are predictor keys. Variable 8, the Off-line Key, correlated negatively with all three criteria and moderately high and negatively with the predictor keys. Variables 9 and 10 resulted in essentially the same pattern since they were constructed from the Off-line Measure.

The Modal Response Key was combined with the Professional Self-confidence Key in multiple predictions of the three criteria to illustrate the effects of the correction scores. The results (presented in Table 10) were as follows: the multiple prediction of the Overall Work Performance Criterion was .15, that for the Creativity Rating, .18, and for the Creativity Check List, .21. Each of these coefficients, although relatively small, show some improvement over the Professional Self-confidence Key when used alone. With higher first order validities for the predictor keys these

Table 10

Maximum Multiple Prediction of the Criteria In the Sample I Analysis*

Predictors	1	2	3
Modal Response Key (Variable 7) and Creativity 2 Key (Variable 8)	.15	.18	.21

*$r_{.05} = .10$; $r_{.01} = .13$

correction measures would have made a more important contribution.

In view of these results, and since the Total Sample Analysis was composed of scientists from six different research centers, it was decided that each center would be studied separately. This kind of analysis was undertaken in order to obtain greater insight into the above relationships, that is, to determine whether the BI would more accurately predict the criteria at some centers than at others and to discover further relationships which would help to more fully understand the predictive power of the BI.

Results of the Sample II Analysis. A total of 269 scientists who had completed the BI and on whom complete criterion information had been obtained were included in the Sample II study.

Five criterion measures were obtained for this analysis. Three of these, Overall Work Performance, Creativity Rating, and the Creativity Check List, are the same criteria used in the Combined Sample Analysis. From the Check List, a control variable called "Likeability" was obtained in order to determine to what extent this kind of factor might influence the supervisor ratings. This likeability variable was constructed from two statements in the Creativity Check List having to do with whether the scientist was pleasant to work with and how well he worked with his supervisor. It was felt that this procedure would be a useful indication of the extent to which this Likeability factor was influencing the supervisors in their performance appraisals, since they were supposed to be considering only creative performances in the completion of the Check List. It was intended that, if common variance was found between this score and the criteria, the Likeability measure would be partialled out or combined with a BI key to more fully determine the dimensions and predictability of the criterion scores.

---

*Although other combinations of the BI keys could have been utilized in multiple prediction, the particular combination of Variable 7 with Variable 6 resulted in the Maximum prediction of the criteria in the Sample I analysis, and therefore is the only combination presented. This samo procedure of key selection for maximum multiple or partial correlation will be utilized in similar tables throughout this report.

Since it was found in the Combined Sample Analysis that the two creativity criteria had a great deal of variance in common but that neither was as predictable as expected, a fifth criterion involving these two measures was constructed which was thought to be somewhat more accurate and predictable. This was accomplished by menas of a special computer program which first converted the Creativity Rating and the Creativity Check List measure for each scientist to a standard score and then computed a mean of the resulting two scores. This procedure was followed for each center separately. The resulting standard scores therefore were based upon each scientist's own group.

The results of the Sample II analysis are presented in Table 11. Variables 1 through 5 are the criteria as described above. As in the Combined Sample Analysis, the first three criteria are characterized by quite high intercorrelations (though not very high if they are considered to be reliability estimates). Variable 4, the Likeability score, correlated with all four of the other criteria to a significant degree. Thus, as expected, this dimension was an influencing factor in the supervisory ratings. This was especially apparent with Variable 1, where the correlation was .37.

Variables 6 through 12 are the BI keys, identical to those described above as used in the Combined Sample Analysis and throughout the Form C-1 study. Variables 6 through 8, the predictor keys, show about the same prediction pattern with the criteria as was shown in the Sample I analysis. As was expected, all three keys were somewhat more successful in their prediction of Variable 5, the combined Criterion, than with either Criterion Variables 2 and 3 taken alone. Variable 9, the Modal Response Key, again seems to be identifying some additional criterion variance other than that identified by the predictor keys. Again Variable 10, the Off-line Key, and its derivatives, Variables 11 and 12, (the False Modesty and Exaggeration Keys) did not perform as expected, since their consideration does not result in increased prediction of the Creativity or Overall Performance Criteria.

Examination of Variable 4 and its relationship with all eleven of the other variables makes apparent a very interesting trend. As was mentioned previously, this Likeability measure correlated significantly positive with all four of the other criteria scores. This variable also correlates positively (although not significantly) with all four of the correction keys, but has slight negative relationships with the three

Table 11

CRITERION INTERCORRELATIONS, KEY SCORE INTERCORRELATIONS, MEANS, STANDARD DEVIATIONS, AND CROSS VALIDITIES FOR FORM C-1 SAMPLE II*

(N = 269)

			1	2	3	4	5	6	7	8	9	10	11	12	Means	S. D.
Criteria	1.	Overall Work Performance	--												9.50	1.75
	2.	Creativity Rating	71	--											8.76	2.09
	3.	Creativity Check List	67	75	--										24.46	8.47
	4.	Likeability	37	20	22	--									18.68	3.81
	5.	Mean Creat. Rating & Ck. List	73	95	93	23	--								50.02	7.63
Predictor Keys	6.	Creativity 1 Key	11	21	20	-11	22	--							98.13	6.97
	7.	Form A Total Score Key	13	22	21	-08	23	88	--						108.77	10.49
	8.	Creativity 2 Key (Prof. Self. Conf.)	10	18	18	-12	20	-74	84	--					100.85	5.36
Correction Keys	9.	Modal Response Key	08	05	08	09	07	-05	18	-06	--				109.71	2.74
	10.	Off-line Key	-04	-11	-11	09	-12	-37	-15	-42	00	--			100.55	3.67
	11.	False Modesty Key	-05	-10	-14	09	-13	-26	-13	-35	-14	81	--		103.63	2.13
	12.	Exaggoration Key	00	04	01	-08	02	39	15	37	16	-78	-39	--	100.00	2.24

Decimal Points omitted. r = .12; r = .16

predictor keys: Those findings give further evidence that, not only was this Likeability characteristic an influencing factor in the supervisory ratings, but it also affected the predictive power of the BI in its relationships with the Creativity and Overall Performance Criteria. These relationships can be further investigated utilizing the procedures for partial and multiple correlation as presented in Table 12.

When the Likeability measure (Variable 4) is held constant, the correlation between the Form A Total Score Key (Variable 7) and the Overall Work Performance Criterion (Variable 1) is raised from .13 to .17 and that between Variable 7 and the Combined Criterion (Variable 5) is raised from .23 to .26. The multiple prediction of Criterion 1 with Variables 4 and 7 is .40 and the multiple prediction of Criterion 5 with these two variables is .34. This latter procedure is not fully justified, since the Likeability score is a part of the other criterion scores and not a predictive measure. A more valid procedure which has not yet been developed would be to construct a BI key to predict Likeability and then determine its relationships with these variables. However, the procedure was carried out here to illustrate the degree of influence which this factor had in the criterion measures and the magnitude of correlation which could be obtained with its use. Use of this technique also has some justification since the Likeability score is overlapping a portion of the criteria (Variables 1 and 5) and the predictor (Variable 7) is overlapping the non-likeability part of the criteria.

It thus seems apparent that this Likeability dimension was indeed one of the difficulties affecting the criterion measures obtained. Although the raters were instructed to consider only the creativity of the ratee, they were obviously influenced by the Likeability dimension, and since

Table 12

Maximum Multiple and Partial Prediction of

The Criteria in the Sample II Analysis

	Partial Correlations	
	Criteria	
Predictors	1	2
Likeability (Variable 4) and Form A Total Score Key (Variable 7)	.17	.26
	Multiple Correlations	
Variable 4 and Variable 7	.40	.34

$r_{.05} = .15$; $r_{.01} = .19$

the BI keys were constructed without taking this into account, this situation resulted in a detrimental effect in the predictability of the criteria with the BI keys. However, a correlation of .40 in predicting Overall Work Performance and .34 in predicting Creativity, which results from these dimensions taken together, is quite a noticeable degree of prediction in light of other potential contaminators and also probable unreliabilities in the criteria.

Results of the Sample III Analysis. Complete data were obtained on a total of 56 scientists who comprised the sample in this study.

The results of the Sample III analysis are presented in Table 13. Variables 1 through 5, the Criterion measures, are identical to those described in the Sample II analysis. The interrelationships among the criteria in this sample (with the exclusion of the Likeability score) are lower than those for the Total Sample or for any other center studied separately. The Likeability score, although not significantly related to Criterion 1, is again observed to be a significant factor in the creativity criteria.

Table 13

CRITERION INTERCORRELATIONS, KEY SCORE INTERCORRELATIONS, MEANS, STANDARD DEVIATIONS, AND CROSS VALIDITIES FOR FORM C-1 SAMPLE III*

(N = 56)

			1	2	3	4	5	6	7	8	9	10	11	12	Means	S. D.
Criteria	1.	Overall Work Performance	--												10.34	3.20
	2.	Creativity Rating	56	--											9.46	2.35
	3.	Creativity Check List	68	67	--										24.82	9.32
	4.	Likeability	18	31	30	--									17.50	5.78
	5.	Mean Creat. Rating & Ck. List	67	92	91	33	--								50.01	4.55
Predictor Keys	6.	Creativity 1 Key	02	-01	05	02	03	--							99.57	7.26
	7.	Form A Total Score Key	02	-05	07	04	02	86	--						108.32	10.33
	8.	Creat. 2 Key (Prof. Self Conf.)	05	10	16	10	15	71	84	--					100.30	5.13
Correction Keys	9.	Modal Response Key	-01	-06	-00	-02	-03	-17	08	-12	--				108.45	2.79
	10.	Off-line Key	-02	-15	-13	-05	-16	-39	-12	-36	-05	--			100.36	3.82
	11.	False Modesty Key	00	-20	-09	-12	-16	-23	-05	-19	-22	78	--		103.84	2.32
	12.	Exaggeration Key	-03	01	04	-01	03	45	17	38	-21	-75	-27	--	95.68	2.52

*Decimal points omitted. $r_{.05} = .26$; $r_{.01} = .34$

The predictor keys, Variables 6 through 8, are again characterized by high interrelationships, but of the fifteen cross validities between these predictors and the five criteria, not one is significantly different from zero. The Correction Keys, Variables 9 through 12, also added no significant variance, generally exhibiting negative correlations with the criteria. It can thus be observed that not only are the criteria characterized by low interrelationships, but nothing else (save Likeability) is related to these criteria either. Needless to say, this sample added little to the effectiveness of the BI in predicting the criteria in the Combined Sample Analysis or to the understanding of the prediction phenomena.

Results of the Sample IV Analysis. The Sample IV analysis was composed of scientists from two different research centers, the two centers being combined because complete data were obtained on only 22 scientists at one of the centers and only 35 at the other; thus, the two were combined to make a sample of 57 scientists.

The results of the Sample IV analysis are presented in Table 14. Variables 1, 2, 3, and 5 (the criteria) are characterized by high intercorrelations, but Variable 4, the Likeability score, was not significantly related to any of these criterin measures. This latter finding is quite significant, since this was the first sample in which such low relationships were found between Likeability and the other criteria. Such a finding indicates that the supervisors in this sample were not significantly influenced by this dimension in their performance apprasials, and thus their ratings should represent a much clearer picture of these scientists' performance than do those ratings obtained at the other centers.

Variables 6 through 8, the predictor keys, were quite successful in their prediction of the criteria in this sample. All twelve of the cross-validities between these predictors and Criteria 1, 2, 3, and 5

Table 14

CRITERION INTERCORRELATIONS, KEY SCORE INTERCORRELATIONS,

MEANS, STANDARD DEVIATIONS, AND CROSS VALIDITIES FOR FORM C-1 SAMPLE IV*

(N = 57)

			1	2	3	4	5	6	7	8	9	10	11	12	Means	S. D.
Criteria	1.	Overall Work Performance	--												9.55	2.27
	2.	Creativity Rating	78	--											8.40	2.23
	3.	Creativity Check List	76	77	--										22.98	8.84
	4.	Likeability	17	14	14	--									17.93	5.56
	5.	Mean Creat. Rating & Ck. List	80	93	94	15	--								50.00	4.39
Predictor Keys	6.	Creativity 1 Key	36	28	32	-09	30	--							101.00	7.84
	7.	Form A Total Score Key	45	35	44	-01	40	90	--						108.68	10.91
	8.	Creat. 2 Key (Prof. Self Conf.)	43	35	44	-03	40	83	89	--					100.53	5.69
Correction Keys	9.	Modal Response Key	13	-03	03	21	00	-18	00	-20	--				108.35	2.34
	10.	Off-line Key	-13	-05	00	07	-02	-56	-31	-49	07	--			99.02	3.91
	11.	False Modesty Key	17	-02	-01	-08	-02	-36	-22	-63	-11	81	--		103.44	2.21
	12.	Exaggeration Key	00	05	-04	-25	-05	46	21	39	-21	-75	-32	--	94.93	2.55

*Decimal points omitted. $r_{.05} = .26$; $r_{.01} = .34$

are significant, nine of them beyond the .01 level. The Overall Work Performance criterion seems to be the most predictable, its correlation with the Total Score Key being .45. The Creativity Check List was also very highly predictable, its correlation with both the Total Score Key and the Creativity 2 Key being .44. The Creativity Rating was also predicted to a magnitude that is significantly different from zero at the .01 level by these latter two keys. The correlations between these three predictor keys and the Likeability score were all negative, though essentially zero; thus, Likeability was apparently not a significant factor influencing the criteria or the BI predictions.

The Correction Keys were also very effective in their performance with this sample. The Modal Response Key is identifying additional criterion variance since it correlates .13 with Variable 1 and -.20 with Variable 8. The False Modesty Key is also effective since its correlation with Variable 1 is .17, with Variable 8 is -.63, and with Variable 9 is -.11.

The partial and multiple prediction of the Overall Work Performance Criterion and a Creativity Criterion (the Creativity Check List), are presented in Table 15. When the Likeability variable is held constant, the prediction of Criterion 1 by the Form A Total Score Key is .46 and the prediction of the Creativity Criterion (Variable 3) by the Creativity 2 Key is .45. The multiple prediction of Criterion 1 with the Total Score Key, the Modal Response Key and the False Modesty Key is .55 and that of Criterion 3 with the Creativity 2 Key and the Off-line Key is .51. These correlations between predictors and criteria are the highest obtained in the entire Form C-1 study and their magnitude indicates a truly remarkable degree of predictive validity for initial job performance. This becomes especially apparent when one realizes the tremendous difficulty involved in the prediction over a period of time of such complex and sometimes subtle criteria.

Table 15

Maximum Multiple and Partial Prediction of the Criteria in the Sample IV Analysis

Predictors	Criteria 1	Criteria 2
Partial Correlations		
Likeability (Variable 4) and Form A Total Score Key (Variable 7)	.46	
Likeability (Variable 4) and Creativity 2 Key (Variable 8)		.45
Multiple Correlations		
Total Score Key (Variable 7) Modal Response Key (Variable 9), and False Modesty Key (Variable 11)	.55	
Creativity 2 Key (Variable 8) and Off-line Key (Variable 10)		.51

$r_{.05} = .32$; $r_{.01} = .39$

Results of the Sample V Analysis. Complete criterion and biographical data were obtained on a total of 92 scientists who comprised this sample.

The results of the Sample V analysis are presented in Table 16. The Overall Work Performance and the Creativity Criteria are again highly intercorrelated, and the Likeability score is again observed to be a significant dimension influencing the criterion measures. None of the fifteen cross validities between the predictor keys (Variables 6 through 8) and the criteria is significantly different from zero. Of the four correction keys (Variables 9 through 12), the Modal Response Key seems to be the only one contributing to the prediction of the criteria. In fact, this correction key is more highly related to each of the criteria

Table 16

CRITERION INTERCORRELATIONS, KEY SCORE INTERCORRELATIONS, MEANS,
STANDARD DEVIATIONS, AND CROSS VALIDITIES FOR FORM C-1 SAMPLE V

(N = 92)

			1	2	3	4	5	6	7	8	9	10	11	12	Means	S.D.
Criteria	1.	Overall Work Performance	--												9.01	2.67
	2.	Creativity Rating	73	--											7.88	2.58
	3.	Creativity Check List	75	83	--										21.25	8.76
	4.	Likeability	34	32	34	--									16.94	6.14
	5.	Mean Creat. Rating & Ck. List	78	96	95	34	--								49.98	10.13
Predictor Keys	6.	Creativity 1 Key	-08	04	03	-07	04	--							101.13	8.16
	7.	Form A Total Score Key	00	11	13	-04	12	87	--						110.19	9.95
	8.	Creativity 2 Key (Prof. Self. Conf.)	05	08	17	-02	13	63	83	--					100.41	4.88
Correction Keys	9.	Modal Response Key	19	16	23	07	20	-12	13	08	--				108.48	2.40
	10.	Off-line Key	-04	02	-06	07	-03	-59	-39	-51	-12	--			100.27	4.05
	11.	False Modesty Key	-08	01	-05	09	-02	-35	-20	-37	-20	85	--		103.58	2.35
	12.	Exaggeration Key	-05	03	06	-09	06	67	50	54	00	-79	-47	--	95.95	2.38

Decimal points omitted. $r_{.05} = .20$; $r_{.01} = .26$

than is any other key in the analysis. Its correlation with Criterion 1 is .19, with Criterion 2, .16, with Criterion 3, .23, and with Criterion 5, .20. These latter two are significant at or beyond the .05 level.

The results of this analysis seems to indicate that among other things we are once again dealing with contaminated and unreliable criteria, since nothing semms to be related to these criterion measures except Likeability and Modal Response. The BI was therefore relatively unsuccessful for the following prediction of criteria in this sample, and like Sample III, this sample added little to the effectiveness of the BI in predicting the criteria in the Total Sample Analysis.

Results of the Sample VI Analysis. This final sample included in the Form C-1 study was comprised of 149 scientists on whom complete data were obtained.

The results of the Sample VI analysis are presented in Table 17. It can be observed that the criteria (Variables 1 through 7) are different from those used in the previous analyses. Variables 1 and 2 are identical to those obtained in the other samples, but the other five are new. The creativity score obtained for Variable 3 came from a longer form of the check list than that used in the previous analyses (this one having 78 items, the former having 51 items). The remaining criterion scores, Variables 4 through 7, all came from this Creativity Check List. Variable 4 was obtained by simply counting the total number of items which a rater had checked for each scientist and including this number as his score. Variable 5 is a ratio score obtained by counting the number of positive items checked and dividing this sum by the total number of items checked. Variables 6 and 7 were obtained as the result of a quite lengthy procedure as described below.

Table 17

CRITERION INTERCORRELATIONS, KEY SCORE INTERCORRELATIONS, MEANS, STANDARD DEVIATIONS, AND CROSS VALIDITIES FOR FORM C-1 SAMPLE VI

(N = 149)

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	Means	S. D.
Criteria		1. Overall Work Performance	--														9.71	3.67
		2. Creativity Rating	75	--													8.66	3.34
	(Creativity Check List)	3. Creativity Check List	71	76	--												19.23	8.42
		4. Total No. Checked	25	31	40	--											58.42	42.72
		5. # Pos. checked/total ck.	70	77	85	26	--										42.02	26.52
		6. Items based on Crit. Val.	61	70	88	32	76	--									21.29	9.13
		7. Items based on Key Score Val.	62	70	69	42	80	56	--								22.99	14.21
Predictor Keys		8. Creativity 1 Key	22	23	21	09	20	30	25	--							97.49	7.29
		9. Form A Total Score Key	20	23	17	05	17	27	22	90	--						106.33	10.55
		10. Creat. 2 Key (Prof.Self.Conf.)	15	19	19	04	19	29	27	77	85	--					99.46	5.65
Correction Keys		11. Modal Response Key	02	-04	-04	-05	-02	-06	-02	-18	00	-21	--				109.06	2.71
		12. Off-Line Key	-01	-02	-15	-10	-12	-19	-21	-43	-19	-40	-03	--			100.67	3.81
		13. False Modesty Key	-01	00	-14	-05	-12	-17	-16	-30	-08	-22	-11	84	--		103.72	2.11
		14. Exaggeration Key	-03	-01	05	05	08	10	18	41	23	44	-08	-76	-41	--	96.16	2.33

Decimal points omitted. $r_{.05} = .16$; $r_{.01} = .21$

The modifications described in the Check List which resulted in Variables 4 and 5 were carried out because it was felt that the Check List was perhaps our best instrument for getting a criterion score on creativity. This consideration further resulted in the derivation of Variables 6 and 7, obtained through a very novel procedure designed to fully maximize the potentialities of the instrument and to provide a more reliable and predictable criterion target. The procedure involved an item alternative analysis of the 78 statement Check List against seven other measures; 2 criterion scores and 5 BI key scores. These were: (1) the Overall Work Performance Criterion; (2) the Creativity Rating Criterion; (3) the Form A Total Score Key; (4) and (5) two forms of the Off-line Key, (6) the Creativity 1 Key, and (7) the Creativity 2 Key (Professional Self-confidence). The results of this analysis thus led to the determination of the relationships which each item in the Check List had with the other seven measures and to determine which items were "live" and which were "dead" weight in respect to each variable. Needless to say, this analysis will be extremely useful in any future work with this criterion form. Through the examination of these results, two scoring keys were developed for the Check List; one based upon the intercorrelations between the Check List statements and the two criterion scores and the other based upon the intercorrelations between the Check List statements and the creativity keys. The former was comprised of 39 statements, sixteen statements having the highest average negative correlations with the two criteria and twenty-three statements having the highest average positive correlations. This key, when applied to the Check Lists obtained for each of the 149 scientists in the sample, resulted in scores which comprise Variable 6. The same procedure was carried out for the second key and resulted in twenty-five statements, twelve negative and thirteen positive, each having the highest average intercorrelations with the two BI creativity keys. This was then applied to each scientist's Check List, the resulting scores comprising Variable 7, which thus is a criterion score built to match a predictor score.

Variables 8 through 10, the Predictor Keys, and Variables 11 through 14, the Correction Keys, are identical to those used in the previous samples.

It can be observed from Table 13 that the first three criteria, Variables 1, 2, and 3 are again characterized by high intercorrelations. Variable 4 shows a significant relationship with Variable 3, the Creativity Check List, so it appears that the sheer number of items checked (both positive and negative) in the Check List does have an influence upon the criterion scores obtained with this measure. However, because of the relationships Variable 4 has with the BI keys, partialling out its effects does not result in an increase in the prediction of the Check List Criterion. It is also apparent from the criterion intercorrelations and cross validities with Variable 5 that little additional information was obtained with its inclusion over what was found from the original Check List Criterion. Thus although the derivation of Variables 4 and 5 did not result in increased prediction of the Creativity Check List Criterion, Variables 6 and 7 reveal significant validity information, as discussed below.

The Predictor Keys, Variables 8 through 10, show significant validation results. Variable 8 exhibits cross validity coefficients which are significant at or beyond the .01 level for each of the three original criterion scores. Variable 9 was just about equally as successful, while Variable 10 was somewhat less so. As mentioned previously, Variable 4 has non-significant relationships with the BI keys, and the cross validities with Variable 5 are of the same magnitude as those with Variable 3 so that little additional information results from the inclusion of these two criterion scores. The other two derivatives of the Check List, Variables 6 and 7, did result in greater predictability of this dimension since the correlation between Variable 6 and Variable 8 is .30 and that between Variable 7 and Variable 8 is .25. The same pattern of increased prediction is apparent with the other two predictors, Variables 10 and 11, where appreciable increases in the cross validity prediction of the Creativity Check List are observed.

The Correction Keys, Variables 11 through 14, generally were unsuccessful since they added little or no additional criterion variance. However, the Off-line Key, Variable 12, was combined with the Creativity 1 Key for maximum prediction of Variables 1 and 2 as observed in Table 18.

Table 18

Maximum Multiple Prediction of the Criteria

in the Sample VI Analysis

	Criteria	
Predictors	1	2
Creativity 1 Key (Variable 8) and Off-line Key (Variable 12)	.24	.25

$r_{.05} = .20$; $r_{.01} = .24$

Summary and Discussion of the Key Score Analysis. It is now apparent that the breakdown of the Total Sample Analysis into separate analyses of the various centers has resulted in greater understanding of the relationships involved. The significant but relatively low validities obtained in the combined sample can now be better understood in light of the results of the various subsample analyses. It will be recalled that about the same results were obtained with Sample II as with the Total Sample. Samples III and V resulted in essentially zero validities while Sample VI resulted in higher validities than in the Total Sample and Sample IV resulted in comparatively high prediction of all the criteria. The task then, is to examine these various results and discover evidence which will help to explain these differences in prediction from center to center.

The most obvious explanation for the low validities obtained in the Total Sample Analysis is the extreme complexity involved in the combined study of 622 scientists from different working environments. This certainly suggests some interesting work climate differences existent in the different research organizations studied, and the further investigation of such differences might lead to more understanding of the results found in the Form C-1 study.

As has been mentioned throughout this report, a major difficulty involved in this complexity lies in the criterion measures obtained from these different research centers. That these measures were not obtained under maximally desireable conditions and the observation that in some centers the criteria was much more predictable than in others (those centers having the smallest number of people to be rated contributing the most predictable criteria) suggests the unreliability and inaccuracy of these pooled supervisory ratings. Further evidence pertaining to this suggestion will be presented in the next section where the item analysis is discussed.

Also in relation to this complexity, it can be observed that for those centers which showed poor validation results, the Likeability dimension was demonstrated to be more of a contaminating influence than for those with good validation results. This feature has been proven to be an especially important factor in the differences found from center to center where the Likeability dimension was observed to be a major influence in the criterion measures obtained at some of the centers and in the prediction of these measures with the BI.

Mention should also be made here concerning the internal consistency of the various BI keys in relation to each other. As expected, in almost every case the predictor keys were negatively related to the correction keys, the latter usually showing some positive relation to the Likeability criterion. With this being the case in the better validation results, these measures were combined in effective multiple and partial correlational methods for maximum prediction of the various criteria.

It is therefore concluded from these considerations that in spite of the tremendous complexity involved, the potential predictive power of the BI has been best indicated in those instances where more adequate criterion targets were available. Reference is made here to the Sample IV analysis, in which cross validities of .55 and .51 were obtained, respectively, in the prediction of the Overall Work Performance and the Creativity Check List criteria.

Item Alternative Analysis in the Form C-1 Study. This analysis was undertaken to determine whether keys specifically constructed for the prediction of the criteria in this study would be more successful in validation results than were the keys developed from previous concurrent validity studies. It was thought that perhaps scientists who are already on the job (as was the case in previous studies) would respond somewhat differently in the BI than those who are applying for a position, which was the case in the present study. If this is so, then the relationships between the BI responses and the criteria would also be affected. Therefore, insofar as the criteria are accurate and reliable, construction of new keys specifically for these responses should provide better validation results.

The procedure involved in this analysis was essentially the same as that described for the item analyses in previous studies. The 622 scientists were split into two Odd-Even samples of 308 and 314, respectively. An item alternative analysis was then carried out in a double cross-validation design by means of a high speed computer at the University of Utah Computer Center. The computer program utilized for this analysis was essentially the same as that used in former studies, except that it has been somewhat revised and expanded.

As a result of this double cross validation design, the correlations reported in Tables 19 and 20 between the various keys and the criteria are cross validities. As these tables indicate, a separate key was developed for each of the four sections of the BI as well as for the total instrument, against each of the three criteria. Section I is concerned with Developmental History, Section II with Parents and Family Life, Section III with Academic Background, and Section IV with Adult Life and Interests.

Table 15 gives the results of the keys developed on the Even Sample as applied to the Odd Sample. The most striking thing about these results is that, for all three of the criterion keys, the only section of the BI contributing validity significantly different from zero is Section IV, the Adult Life and Interest Section. In each case, the validity coefficients exhibited by any of the other sections is essentially zero, thereby leaving Section IV the only valid part of each key. Thus, the correlation between Variable 7 and Variable 1 is .14, between Variable 12 and Variable 2, .18, and between Variable 17 and Variable 3, .25. It will be noted that these cross validities, although rather low, are all significant at or beyond the .01 level of confidence.

Another interesting feature of these results is that Variable 17 exhibited higher cross validities (.22 and .26) for the other two criteria (Variables 1 and 2) than any of the keys specifically constructed for the prediction of these two criteria.

Table 19

CRITERION INTERCORRELATIONS, CROSS VALIDITIES, MEANS AND

STANDARD DEVIATIONS FOR FORM C-1 ITEM ANALYSIS - ODD SAMPLE

(N = 308)

		1	2	3	Means	S. D.
Criteria	1. Overall Work Performance	--			9.57	1.98
	2. Creativity Rating	68	--		8.70	2.14
	3. Creativity Check List	67	72	--	49.60	10.24
Even Sample Keys	4. Section I Overall Performance Key	06	-01	-04		
	5. Section II " " "	04	-07	-01		
	6. Section III " " "	-10	-04	-05		
	7. Section IV " " "	14	10	15		
	8. Total Score " " "	08	-02	01	97.19	2.08
	9. Section I Creativity Rating Key	07	06	-04		
	10. Section II " " "	04	-05	-02		
	11. Section III " " "	09	08	05		
	12. Section IV " " "	15	18	20		
	13. Total Score " " "	18	16	11	99.39	2.01
	14. Section I Creativity Check List Key	07	08	-03		
	15. Section II " " " "	-00	-07	01		
	16. Section III " " " "	02	03	04		
	17. Section IV " " " "	22	26	25		
	18. Total Score " " " "	20	22	18	98.45	3.00

Decimal points omitted. $r_{.05} = .11$; $r_{.01} = .14$

Table 20 gives the results of the keys developed on the Odd Sample as Applied to the Even Sample. Although the criteria intercorrelations are somewhat higher for this sample, these results show about the same pattern as those for the Odd Sample, Section IV exhibiting the highest correlations in each key. Variable 7, Section IV of the Overall Performance Key, exhibited the highest correlations with all three criteria, higher than any of the keys specifically constructed for the prediction of the other criteria. Thus, the maximum prediction of the criteria as shown with this key are .15 for Variable 1, .17 for Variable 2, and .21 for Variable 3.

The superior performance of Section IV in this analysis is entirely consistent with all of the former studies of the Biographical Inventory, where this section resulted in better prediction of the criteria than the other parts of the instrument. The difference between results in this and in the former studies lies in the rather low validities of all the BI sections, and especially the extremely low and sometimes negative validities of Sections I, II, and III. With this low magnitude for the first three sections, Section IV actually predicted the criteria to a higher degree than did total scores across all four sections.

Summary and Discussion of the Item Analysis. It can thus be observed that when considering only the case of the Total Sample, the item analysis of the Form C-1 data resulted in somewhat better prediction of the criteria than had the application of previously constructed keys to these same criteria, although the increase was rather small and quite variable from the Odd to the Even Sample. Thus the maximum prediction of the Overall Work Performance Criterion in the Combined (Sample I) Key Score analysis was .15 while in the item analysis it was .19 (this latter figure obtained by averaging the highest correlation between a key and this criterion in each of the Odd-Even Samples). For the Creativity Rating

Table 20

CRITERION INTERCORRELATIONS, CROSS VALIDITIES, MEANS, AND STANDARD DEVIATIONS FOR FORM C-1 ITEM ANALYSIS-EVEN SAMPLE

(N = 314)

		1	2	3	Means	S.D.
Criteria	1. Overall Work Performance	--			9.54	2.14
	2. Creativity Rating	74	--		8.51	2.41
	3. Creativity Check List	72	77	--	49.38	10.28
Odd-Sample Key	4. Section I Overall Perf. Key	01	03	03		
	5. Section II " " "	09	06	06		
	6. Section III " " "	01	04	05		
	7. Section IV " " "	15	17	21		
	8. Total Score " " "	11	15	18	98.73	2.80
	9. Section I Creativity Rating Key	-02	01	04		
	10. Section II " " "	00	02	-06		
	11. Section III " " "	-02	-01	-00		
	12. Section IV " " "	11	14	20		
	13. Total Score " " "	08	12	16	100.58	3.16
	14. Section I Creativity Ck. List Key	00	06	04		
	15. Section II " " " "	-03	-01	-04		
	16. Section III " " " "	-03	05	03		
	17. Section IV " " " "	08	11	17		
	18. Total Score " " " "	04	12	14	100.56	2.61

Decimal points are omitted. $r_{.05} = .11$; $r_{.01} = .14$

Criterion, the key score analysis resulted in a prediction of .18, and for the item analysis, .22, and finally, for the Creativity Check List Criterion, .21 and .23 respectively, for the two types of analyses.

A methodological problem needs to be mentioned here. Because of the relative unreliability of the criteria and other measurement contaminants in this study, the particular type of biserial item analysis procedure utilized in this analysis resulted in only an average of 6 or 7 item alternatives being scored for each subject in the prediction of each criterion. When the shortness of these keys is taken into account, a cross validity coefficient of .26 in predicting creativity is quite remarkable. However, had more items been utilized in these predictions, the validities might have been increased.

These results provide evidence in relation to one of the main questions under consideration in this analysis. That is, construction of new keys specifically designed to predict the criteria in this sample did not result in appreciable increases in prediction over what had been obtained with the application of previously constructed keys to these same criteria. This indicates further evidence in support of the contention that the criteria under consideration in this study are not adequate. This statement holds true for the total sample considered together, although more information could be obtained relative to these relationships if each center were subjected to separate item analyses. However, the statistical procedures involved depend upon a large number of subjects, which would not be the case if each center were studied separately. Therefore, based upon the evidence now available, some stability and reliability of these previously constructed keys has been demonstrated, even in the prediction of questionable criteria.

Concluding Remarks. The Biographical Inventory has now been studied under several concurrent validity settings, and, with the completion of the Form C-1 study, it has also now been studied under predictive (longitudinal) conditions. The proven effectiveness of the instrument in the former concurrent situations has held up quite well for some of the centers (Samples IV and VI) in the longitudinal design, and not so well for some of the others (Samples II, III, and V). Several possible explanations have been entertained in this report to account for these differing results. The following discussion brings several other relative considerations to light.

The extreme complexity in such a study as reported here cannot be overemphasized. Studying so many scientists from different working environments and from different areas of interest introduce an especial strain upon the criterion problem, especially when pooling samples togehter across NASA centers. It has been shown in the Form C breakdown analysis that prediction increases when the subjects under investigation are working in more homogenous areas of endeavor. The same has been shown in the Form C-1 study where increased prediction resulted in some cases when the different centers were studied separately. However, this is only one of the problems involved in this complexity; many others must remain unaccounted for.

Another aspect of this problem is that there are undoubtedly many differences in the group of scientists involved in the Form C-1 study and those who were involved in the previous investigations. One of these differences is the educational status of these groups. In the Form A study of the BI, 24 per cent of the scientists had a Bachelor's degree or less; 38 per cent had some graduate work, but no degree; 13 per cent had a Master's degree, 16 per cent had some graduate work beyond the Master's degree; and

9 per cent had the Ph.D. degree. The corresponding percentages in the Form C-1 study were as follows: 68 per cent had a Bachelor's degree or less; 23 per cent had some graduate work, but no degree; 6 per cent had a Master's degree' 2 per cent had some graduate work beyond the Master's degree; and 1 per cent had the Ph.D. degree.

Although these and other considerations must be taken into account, it is here posited that if more adequate criterion data were obtained for the scientists participating in the Form C-1 study even at this time, which is now some two years after the BI's were completed, more satisfactory validation results would be obtained. This statement is especially supported by several of the findings reported in this study, where it was shown that criterion data obtained after only one year of observation was less reliable and relevant than that obtained in the previous concurrant validity studies where the observations were based upon a much longer period of time. Nevertheless, it is here concluded that the BI has been demonstrated to be a valuable predictive instrument considering the difficulty and importance of the predictions made, and could contribute valuable information to the identification of scientific talent in a carefully designed selection program.

Chapter 6

Examples of Item Content of the Biographical Inventory

In this section a few examples of some of the better biographical items will be presented with a brief discussion of the types of items which have generally failed to contribute to the identification of scientific talent across the studies completed to date. It should be remembered that the following relationships are characteristic only of the majority; there would be some individuals whose responses to each item would be exceptions to the general finding. All of the items cited were related to the Creativity Criterion and on occasions to other criteria as well. Since there may be some distortion in the responses of the subjects, the extent to which the subjects' responses correspond to the actual situation described remains to be determined. This is an important question to which further research should be directed. Results by Cline, Richards, Abe, and Needham (1963) indicate that biographical items completed by high school students do not describe the home environment in the same way as the parents see it. This, of course, does not say whose perception is correct.

A number of items demonstrated that characteristics of self-determination and an individualistic orientation (or inner-directedness) are positively related to the criteria. A facet of this is concerned with how the individual scientist elects to expend his energies and to what area of his life he devotes himself. For example, a definite task-orientation appears to be involved in the following question. If an individual responds that, to a great extent, he is the kind of person who becomes so absorbed in his work and interests that he does not mind a lack of friends, this response was positively related to the criteria, whereas another person's response that this does not describe him at all was negatively related to the criteria. Another example of an item in this area is as follows: "Assume you are in a situation in which the following two [illegible] alternative courses of action arise. Which one of the two would you be most likely to do? (A) Be a good team man so that others like to work with me, or (B) Gain a reputation through controversy, if necessary, as one whose scientific word can be trusted." Response A was correlated negatively with the criteria, and Response B was positively related. Wherever this attitude of independence originated it evidently tended to have been present during the student's academic career. For example, if the scientist reported that he questioned his professors on subject matter considerably more often than average, his response was positively related to the criteria.

The relationship of undergraduate college grade-point average to success as a scientist has been shown by many investigations to at best be low; however, occasionally a few items in the Academic Section of the Biographical Inventory which are concerned with self-reported academic performance emerge with a low but consistent relationship to the Creativity Criterion. For example, a B.A. or B. S. degree or less has a negative relationship to the criteria, whereas obtaining the Ph.D. degree has a positive relationship. If a scientist described his college undergraduate work as being well above average and himself as being satisfied with his progress, this response was positively related to the criteria. If a scientist reported that as a student he succeeded exceptionally well in his engineering courses or biological science courses, this had a positive relationship to the criteria, while a response of succeeding fairly well had a negative relationship. Other items, such as those concerned with success in the relevant appearing fields of mathematics, physics, and chemistry, have not consistently shown a relationship with the Creativity Criterion.

One of the more consistently surprising items which has demonstrated a positive relationship to creativity is an item which is concerned with attitudes toward making repairs around the house prior to the age of 18. If the subject responds that he had a strong dislike for making such repairs, this response was positively related to creative performance. It is suspected that this item is related to the personality factor of femininity. Previous research has shown that this dimension has some relationship to creativity. It may also reflect certain sensitivities and an orientation toward ideas and theoretical approaches as opposed to more tangible and mechanical interests.

This discussion would not be complete without a brief statement of the types of items that have failed to discriminate. Generally speaking, items that measure a small specific segment of previous experience or a specific fact in one's life history have not been fruitful. For example, items, such as the extent of participation in childhood job enterprises such as cutting lawns, washing cars, etc., or the number of times that the subject had changed residences by the time he entered college, or the age at which he held his first paying job, or the highest level of achievement he obtained in the Boy Scouts, have not survived the validation process. Another area which has so far proved barren for identifying scientific talent concerns descriptions of various parental characteristics, such as the parents' dominance, affection, encouragement, strictness, permissiveness, etc. While it is expected that this is an area of definite importance, it has proved to be extremely difficult to cultivate successfully. One of the reasons for this is probably the complex network of interactions that exists between the subject's parents, so that when any one facet of their behavior has been measured it does not provide enough information about how the other parental characteristics interact; thus, by itself the parental characteristic being measured appears unimportant.

It would be difficult to estimate the number of items which have either been tried out in one form of our inventory or have been carefully examined for their potential discriminating power. Certainly the number exceeds 1,000 items. Undoubtedly it would be possible to construct additional valid items to add to the Biographical Inventory, but according to our current understanding and measurement skills most of the fertile ground has already been plowed. Consequently, gains in the near future through item construction will probably be small, although not necessarily unimportant.

Chapter 7

Other Research Activities with Biographical Information

Identifying Scientific Talent with Biographical Information in "Other" Settings and Samples. Because of the consistently promising and positive results in studies where biographical information was used in developing predictive keys (such as predicting creativity among NASA scientists, etc.) there has been an increasing use of psychological instruments of this type in a broad variety of settings. One of the direct outcomes of this has been the initiation of a research project, based upon this previous NASA effort by Richardson-Merrell, Inc., an industrial pharmaceutical company that is concerned with the development of a biographical inventory for the identification of scientific talent, especially in recruiting new personnel. Just completed in this study (Cline and Tucker, 1965), a biographical inventory (constructed by C. W. Taylor and R. L. Ellison for research use in industry) was administered to a large number of scientists and criterion information was obtained having to do with creativity and general scientific competence. Each participating scientist was rated by his supervisor, his peers, and in some cases his subordinates. Results of this study showed that the empirical keys, constructed during this study, obtained high initial validities, the correlations ranging in the .70's and .80's and cross validity predictions of the criteria in the .30's and .40's. Also, since the biographical information items used in this study contained many of the same items that were also used in the NASA studies, it was possible to score the test protocols of the pharmaceutical scientists with the NASA derived keys.[5] The results of this procedure proved definitely significant and have far ranging implications and application to this entire area of research. It was found that the NASA Creativity Key (Form A, Total Score Key), when applied to the biographical information responses of the Richardson-Merrell pharmaceutical scientists, cross validated .35 with the Creativity Rating Criterion of these scientists. The NASA Off-Line Key (a correction key) cross validated .10 with the Creativity Criterion, and these two keys had a correlation of -.22 with each other. When these two keys were taken together in a multiple correlation with the Creativity Criterion, a multiple cross-prediction correlation of .39 resulted. This finding indicates a significant and impressive applicability of the results of the NASA studies and the wide generalizability of the NASA keys across very diverse samples of scientists (in this case from the physical to biological sciences and from a government agency to private industry). This study therefore suggests a high potential value of biographical information in the identification of scientific talent in a broad variety of organizational settings.

---

[5] It might be mentioned that strict security has been maintained in using these keys for this research purpose only.

Chambers (1964) studied personality and biographical factors of mature scientists who are highly creative in research work and those who are much less creative. Along with significant differences between creative scientists and their controls on several personality variables, he also found significant differences for 16 biographical items. He was thus able to present a biographical and personality profile of those highly creative and not so creative scientists whom he studied.

W. A. Owens and his associates (Kulberg & Owens, 1960; Morrison, Owens, Glennon, and Albright, 1952) have been involved in several studies of engineers and scientists with biographical data in relation to various aspects of creativity, professional interests, and research competence. In a related study, Albright and Glennon (1961) found that biographical information could discriminate between supervisory and research oriented scientists at all levels of a laboratory organization. Also, Smith, Albright, and Glennon (1961) demonstrated the value of the personal history technique in the prediction of criteria of scientific competence and creativity with a highly select group of research scientists.

In a recent study by McDermid (1965) with the technical and engineering personnel of the Hammond Organ Company, it was found that of the several predictive instruments used, only biographical data proved to be significant as predictors of both supervisory and peer ratings of creative performance.

The June 1965 national research conference on the biographical inventory approach chaired by E. R. Henry and supported by the Richardson Foundation, discovered from practically all 16 participants that across professional and other high level, complex fields, the biographical inventory is at least as good and is usually better than other devices in predicting job performance. Also, a biographical and multiple criterion study has been tentatively planned by McPherson and Cooley with the Dow Chemical Corporation personnel. Thus, it appears that the potential value and promise of biographical information is now being recognized in many studies across very diverse criterion groups.

Heterogeneity of the Biographical Inventory. Recently a small piece of exploratory research was completed which highlights the complexity and the stability of the Biographical Inventory. In Form C, 12 landmark personality items were selected from French (1953) in order to relate the biographical items to previous research findings on typical personality instruments. These landmark factor items were as follows: Dominance, Interest in Philosophy, Masculinity-femininity, Persistence, Self-confidence, Sensitive Attitude, Sociability, Gregariousness, Emotionality, Autistic Thought, and Intelligence. In an exploratory attempt to see how the biographical items related to these dimensions of personality, six of the landmark personality items were held up as criterion scores in an item analysis. Although the landmark items themselves had a very low correlation with the regular criterion measures of Creativity, Productivity, etc., scores from the biographical keys which were built to predict these personality landmark items correlated substantially higher with the regular criterion measures. For example, the item measuring Dominance correlated .12 with Creativity; yet the Dominance score from the biographical key of approximately 74 items correlated .37 with the Creativity Criterion. Thus, even though the Dominance Criterion item had a very low relationship with the Creativity Criterion, the key built to predict the Dominance item also succeeded in predicting the Creativity Criterion. This was to be expected, at least to some extent, because of the much greater length (74 items versus 1 item) and thus the consequent greater reliability. Another probable reason is that many of the items which went into the lengthened key were complex factorially and thus measured other aspects of the Creativity Criterion. This is supported by the fact that scores from the Dominance Key correlated .83 with the biographical scores that were specially keyed for the Creativity Criterion. The cross-validity coefficient of this same empirically keyed Dominance score against the Dominance Criterion item was .50. Of the other five landmark items which were held up as criteria in an item analysis, all were predicted with cross validities ranging from .38 to .62, the majority being in the .50's. These findings again illustrate the complexity of the Inventory as it predicted these different areas of personality. This procedure of keying biographical items against a landmark personality item provides an efficient way of building a longer test for that personality characteristic.

Nielsen (1963) has recently found that a biographical inventory which he factored was made up of some twenty to thirty relatively independent dimensions. This again is impressive evidenceof the diversity and complexity of the inventory and why it is able to predict such a variety of criteria, including complex ones.

Abe (1963) has recently completed a study which is relevant in this area. In Abe's study one of the correlation matrices from the Form A study discussed previously was factor analyzed. This matrix was made up of forty-seven variables including seven criteria. The other forty variables were biographical keyed scores. Abe found nine factors which ranged from a pure criterion factor to an outer-directedness factor which seemed to reflect a desire on the part of some scientists for recognition and a need for visibility. This study again illustrates the variety of interesting useful information which can be obtained through research with biographical information.

Biographical Information and the Identification of Scientific Talent in Younger Age Groups. A question has been raised repeatedly that the items in the Biographical Inventory are primarily concerned with topics which are pertinent only to adults and to actual on-going research activities. This question implies that these biographical items would not be very appropriate or valid if administered to younger age groups, such as high school students or college seniors, because the items are oriented too much toward the activities of mature research scientists.

A recent study on this problem supported by the National Science Foundation (C. W. Taylor, Cooley, and Nielson, 1963) highlights some of the complexities involved in early identification of scientific talent, since it implies that our present educational program is not geared to give the most appropriate kind of training as far as creative scientific achievement is concerned. In the NSF-supported summer science program for high school students, some of the students have the unusual opportunity to participate full time in research activities. Others participate in classroom activities only to learn important science materials not in the regular curriculum. The main interest of this high school study, was to determine whether the creative and productive characteristics found for scientists on the job, as discovered in recent studies of Air Force, Space Agency, and other scientists, are measurable on high school students in these programs and whether these same characteristics are more related to the performance of the students in these research activities than to classroom-only performances.

The data analysis has revealed that two distinct groups can be identified, a research achievement group and an academic achievement group. In general the predictors with positive validities for the academic program tended to have low, zero, or negative validities in the research programs,and vice versa. In this study the Biographical Inventory was modified to be appropriate for younger age groups. It was found that the vast majority of the items could be used without modification; a few were revised, and a few were dropped. Because some items had to be rewritten, the scoring keys were constructed on mature scientists and predictive (short-range follow-up) validities rather than concurrent validities were to be determined, it was expected that the revised Biographical Inventory would not work very well, if at all, under the circumstances. The results indicated that of all the instruments used in this study, the Biographical Inventory was the best overall predictor of creative performance. In one of the research participation groups in which it was felt that the most valid criteria was obtained the Biographical Inventory scores correlated .47 with supervisory ratings on creativity. Needless to say, this is a remarkably satisfactory cross validation finding. Certain biographical keys that worked well in the research programs did not work well for the academic programs and vice versa. The two extreme examples are that the scores from the Professional Self-confidence Key were valid for two-thirds of the criteria in the research programs but had no significant validities whatsoever in the academic sample whereas scores from the Miscellaneous Biographical Key were as good as any biographical scores in the academic programs but had no significant validities in the research programs.

It was thought that the Biographical Inventories constructed for administration to NASA scientists would probably be more appropriate for college seniors than for high school students, since college seniors more closely resemble the adult samples upon which the Biographical Inventory was developed. Some data have been obtained on this latter issue already, although the criteria were not as directly pertinent as those used in the study of high school students. Victor Bunderson has been working on an evaluation of present and potential fellowship selection information at the University of Utah, where the Research Committee awards approximately 40 graduate fellowships per year. Selection has been based largely on grade-point average and open-ended letters of recommendation. This study seeks to evaluate these sources of information for fellowship selection in thelight of various criteria of graduate student performance. As a part of this research, a modified Biographical Inventory was administered for research purposes only to a nmmber of seniors and graduate studen who applied for a fellowship. Ratings were obtained a year later on their graduate student performances, including their research potential. Again the Biographical Inventory scores proved to be the most valid predictor of these multiple criteria; in fact, early indications are that the Biographical Inventory, by itself, overshadows the validity of the official, collective judgment of the Fellowship Committee who had used the entire folder of materials for each applicant in making their fellowship decisions.

Future Biographical Research Activities. An examination of the different types of items included in the various Biographical Inventories shows that in the number of characteristics measured, they are very heterogeneous and complex. One of the activities recently completed is an intercorrelation and factor analysis of the biographical items, along with appropriate criterion scores. Such an analysis will yield a great deal of information about this type of invent

Of special interest is the possibility that factor analysis will contribute to the development of more independent and efficient subscores within the inventory than our existing subjective classification of Developmental History, Parents and Family Life, etc., has yielded. This in turn should contribute higher validity coefficients derived from combined subscores, thus increasing the predictive potential of the Biographical Inventory. Another type of useful information from a factor analysis of the items will be the identification of the most promising areas in the inventories. From these leads, it should be possible to construct new items and thereby further improve the instrument.

The Biographical Inventory has usually been found to measure somewhat different criterion variance than other traditional types of selection tests. It therefore seems advisable to consider research on some of these other kinds of measures found to have promise in research on creative scientific talent to see how well they supplement the Biographical Inventory scores. Such additional validation work could take full advantage of the criterion data as well as the biographical data already available on NASA scientists.

In summary, all our research results obtained to date indicate that biographical information is a very promising, if not the most promising, single means of identifying creative and other types of scientific talent.

The cross-validity coefficients obtained are considerably higher than those typically reported for the prediction and identification of creative or of other types of scientific talent, which use other kinds of predictors such as high-level aptitude tests, intelligence measures, college grade-point averages, and personality test measures. It is our conviction that continued research should be carried out to exploit thoroughly the potential in the biographical approach so the identification of creative scientific talent can be accomplished with as much accuracy as possible.

REFERENCES

Abe, C. A factor analytic study of biographical predictors of scientific potential. Unpublished M. A. thesis, University of Utah, 1963.

Adjutant General's Office, Personnel Research Section, War Department. Validation of a program for the selection of officers for retention in the peace-time army. Report No. 703, 1946.

Albright, L. E., and Glennon, J. R. Personal history correlates of physical scientists' career aspirations. J. appl. Psychol., 1961, 45, 281-284.

Alf, E. D., Jr. Comparison of predictive and concurrent validities of basic test battery test scores. Technical Bulletin 63-12. Bureau of Naval Personnel, Dept. of Navy, Washington, D. C., October, 1963.

Bittner, R. H. Quantitative predictions from qualitative data: predicting college achievement from biographical information. J. of Psychol., 1945, 19, 97-108.

Bunderson, C. V., Rigby, Janice, and Taylor, C. W. The validity of fellowship selection information. Technical Report No. 1, University of Utah, 1963. (Multilithed)

Business Week. Good scholars not always best. McGraw-Hill: N. Y., Feb. 24, 1962, 77-78.

Chambers, J. A. Creative scientists of today. Science, 1964, 145, 1203-1205.

Cline, V. B., & Tucker, M. F. The prediction of creativity and other performance measures among pharmaceutical scientists, 1965. (Mimeographed)

Cline, V. B., Richards, J. M., Jr., Abe, C., and Needham, W. Use of biographical information on students who do and do not achieve in high school science courses. Off. of Educ. Coop. Research Project No. 1107, University of Utah, 1963.

Ellison, R. L. The relationship of certain biographical information to success in science. Unpublished M. A. thesis, University of Utah, 1960.

Ellison, R. L., and Taylor, C. W. The development and cross-validation of a biographical inventory for predicting success in science. Paper read at American Psychological Association, St. Louis, 1962.

Ellison, R. L. The development of scoring procedures for a biographical inventory to predict various criteria of success in science. Unpublished doctoral dissertation. Univer. of Utah, 1964.

French, J. W. The description of personality in terms of rotated factors. Princeton, N. J., Educational Testing Service, 1953.

Ghiselin, B. Ultimate criteria for two levels of creativity. In C. W. Taylor and F. Barron (Eds.) *Scientific creativity: its recognition and development.* Wiley: New York, 1963, 30-43.

Golovin, N. E. The creative person in science. In C. W. Taylor and F. Barron (Eds.) *Scientific creativity: its recognition and development.* Wiley: New York, 1963, 7-23.

Guilford, J. P. *Fundamental statistics in psychology and education.* New York: McGraw-Hill, 1950.

Guilford, J. P. *Fundamental statistics in psychology and education.* New York: McGraw-Hill, 1956a.

Guilford, J. P. The relation of intellectual factors to creative thinking in science. In C. W. Taylor (Ed.) *The 1955 University of Utah research conference on the identification of creative scientific talent.* Salt Lake City: Univer. of Utah Press, 1956.

Hansen, P. J. The relationship of certain biographical information to the academic and behavioral adjustment of high school boys. Unpublished doctoral dissertation, Univer. of Utah, 1950.

Kornhauser, A. W., and McMurry, R. N. A scientific procedure for the selection of salesmen. *J. appl. Psychol.*, 1941, *24*, 11-17.

Kulberg, G. E., and Owens, W. A. Some life history antecedents of engineering interests. *J. educ. Psychol.*, 1960, *51*, 26-31.

Lacklon,R., and Harmon, L. Criterion committee report. In C. W. Taylor (Ed.) The second (1957) University of Utah research conference on the identification of creative scientific talent. Salt Lake City: Univer. of Utah Press, 1958.

McDermid, C. D. Some correlates of creativity in engineering personnel. *J. appl. Psychol.*, 1965, *49* (1), 14-19.

Meltzer, L., and Salter, J. Organizational structure and the performance and job satisfaction of physiologists. *Amer. soc. Rev.*, 1963, *27*, 3.

Morrison, R. F., Owens, W. A., Glennon, J. R., and Albright, L. E. Factored life history antecedents of research performance. *J. appl. Psychol.*, 1962, *46* (4), 281-284.

Nielsen, E. C. The use of biographical information in predicting hospital aide performance. Unpublished masters thesis, University of Utah, 1963.

Roe, Anne. A psychological study of eminent biologists. *Psychol. Monogr.*, 1951, *65* (14) Whole No. 331.

Roe, Anne. A psychological study of eminent psychologists and anthropologists, and a comparison with biological and physical scientists. *Psychol. Monogr.*, 1953, *67*, (2) (Whole No. 287).

Roe, Anne. Personal problems and science. In C. W. Taylor and F. Barron (Eds.) *Scientific creativity: its recognition and development*. Wiley: New York, 1963, 132-138.

Schaeffer, E. S., and Bell, R. Q. Development of a parental attitude research instrument. *Child Develop.*, 1958, *29*, 339-361.

Sessions, F. Q. & Taylor, C. W. In-service versus out-of-service administration of tests and criteria. *Personnel Psychol.*, 1961, *14* (1), 67-73.

Smith, W. J., Albright, L. E., and Glennon, J. R. The prediction of research competence and creativity from personal history. *J. appl. Psychol.*, 1961, *45*, 59-62.

Sorenson, G. Non-intelbctual factors related to success in college. Unpublished doctoral dissertation, Univer. of Utah, 1950.

Sprecher, T. B. An investigation of criteria for creativity in engineers. Unpublished doctoral dissertation, Univer. of Maryland, 1957.

Taylor, C. W. (Ed.). *The 1955 University of Utah research conference on the identification of creative scientific talent*. Salt Lake City: Univer. of Utah Press, 1956.

Taylor, C. W. (Ed.). *The second (1957) University of Utah research conference on the identification of creative scientific talent*. Salt Lake City: Univer. of Utah Press, 1958.

Taylor, C. W. (Ed.). *The third (1959) University of Utah research conference on the identification of creative scientific talent*. Salt Lake City: Univer. of Utah Press, 1959.

Taylor, C. W. (Ed.). *Creativity: progress and potential*. McGraw-Hill, New York, 1964.

Taylor, C. W. (Ed.). *Widening horizons in creativity*. Wiley: New York, 1964.

Taylor, C. W., and Barron, F. (Eds.). *Scientific creativity: its recognition and development*. Wiley: New York, 1963.

Taylor, C. W., Cooley, G. N., and Nielson, E. C. Identifying high school students with characteristics needed in research work. NSF-G17543, Univer. of Utah, 1963. (Mimeographed).

Taylor, C. W., Smith, W. R., Ghiselin, B., and Ellison, R. L. Explorations in the measurement and prediction of contributions of one sample of scientists. USAF, AFSC, Personnel Lab., Lackland AFB, ASD-TR-61-96, 1961.

Taylor, D. W. Variables related to creativity and productivity among men in two research laboratories, In C. W. Taylor (Ed.) *The second (1957) University of Utah research conference on the identification of creative scientific talent*. Salt Lake City: Univer. of Utah Press, 1958.

APPENDIX

RATING SCALES FOR NASA SCIENTISTS
For Research Purposes Only

Introduction

The attached rating form is for the purpose of evaluating the individual scientists who are completing the Biographical Information Inventory. These ratings will be kept strictly confidential and will be used for research purposes only. The data will be processed entirely by the research team at the University of Utah, and the results will be reported to NASA only in terms of relationships found on the total group, without any reference whatever to individual scientists or their names.

Instructions

On the attached forms four traits are described, one to a page. Down the left hand side of the page are listed the names of the men you are to rate. On the right hand side of each page are listed the descriptions of seven different degrees of a trait. These descriptions are presented at every other point along a 15 point scale. When rating a man, carefully consider all of the seven descriptive statements. Read all the statements thoroughly before rating a man. Then write the number of the statement which best describes the man in the space provided by his name. You may rightly think that no one of the statements is an exact description of the person but make the best choice you can. The odd numbers between the statements indicate a position on the scale between the two statements. Use one of these numbers if you think it describes the man's work more accurately.

When you have rated every man on that particular trait, proceed to the next page and repeat the process. In rating a person, consider the person's entire day-by-day performance as well as his best performance. In rating the men, remember that everyone has certain strengths and weaknesses and thus it would be a rare instance when a man is equally strong or weak at all aspects of his job. Try to give a comprehensive and, above all, an ACCURATE picture of how your men stand in relation to each other and to the other scientists you have known.

The percentages down the left hand side represent an hypothetical breakdown of how 100 typical scientists might be distributed across the descriptive statements. These percentages are only included as guidelines for your information.

Your cooperation is greatly appreciated.

Write the number
from the scale
which best des-
cribes each man
in the parenthesis
beside his name.

Consider the total quantity of work produced compared to the amount expected in the type of position and at the level of the person being evaluated. Think of his speed of work and dependability in producing and in getting things done. DO NOT consider other aspects of his performance--ONLY the quantity of work he gets done.

( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )

Hypothetical breakdown of how a sample of 100 scientists might be distributed across the descriptive statements.

5% 10% 20% 30% 20% 10% 5%

1.

2. — His output of work is definitely below the amount expected of a person in his position. He is slow and inclined to be careless and superficial.

3.

4. — The way he handles assigned problems is a little disappointing, and he must be encouraged to keep his work moving. He is about as likely to get a job done late as on time.

5.

6. — He is on the borderline, capable of fairly good or average work but inclined to fall below that level. He is occasionally late in getting his work done.

7.

8. — He produces an average amount of work, usually gets his work done on time, and like most of us, occasionally makes a mistake in not doing part of his work.

9.

10. — He produces somewhat more than the average, can be counted upon to get the job done, and in a dependable manner. He makes a real effort to do a good job.

11.

12. — He consistently turns out more work than the average. His hard work, efficiency, and productivity have earned him the respect of his associates.

13.

14. — He produces an outstanding amount of work, is an exceptionally fast and efficient worker, and is also a very accurate one.

15.

(WHEN YOU HAVE RATED EACH MAN, TURN THE PAGE.)

Write the number
from the scale
which best des-
cribes each man
in the parenthesis
beside his name.

Rating Scale B

Consider the man's tact, courtesy, friendliness, patience, and his power to promote group activity and solidarity. Rate the extent to which he senses the needs of others, and his willingness and effectiveness in sharing his knowledge with others. DO NOT consider other aspects of his performance--ONLY his skill in getting along with people.

( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )

Hypothetical breakdown of how a sample of 100 scientists might be distributed across the descriptive statements.

5%
10%
20%
30%
20%
10%
5%

1.

2. — His inability to get along with people is such that it would often be better to assign him to individual work. He has little tact and very few close friends.

3.

4. — Although he gets along with his colleagues, few of them like him well. He works better by himself and he is not very effective in presenting his ideas or results to others.

5.

6. — He is liked by many members of his group and he can work effectively with the right people. Ordinarily, however, he would not be chosen to represent his group in outside contacts.

7.

8. — He is liked by most of those in close contact with him and he works satisfactorily as a team member. He is reasonably effective in dealing with other individuals or groups from outside his laboratory.

9.

10. — A good team member, often being successful in getting others to work well together. He makes a good impression, expresses himself well, and occasionally represents his group in outside contacts.

11.

12. — He makes friends easily and has considerable skill in handling people. Almost always successful in getting effective cooperation among members of a group. Is particularly effective in presenting his ideas.

13.

14. — He is well liked by nearly everyone. His skill in getting others to work together effectively, in handling difficult interpersonal situations, and in representing his group is exceptional.

15.

(WHEN YOU HAVE RATED EACH MAN, TURN THE PAGE.)

Write the number from the scale which best describes each man in the parenthesis beside his name.

( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )

Rate the product of the man's work as to its creativity. Consider the implications of his work, its impact, the originality of the approaches used by the scientist, the comprehensiveness and novelty of the solutions, the degree to which his work has opened the way and stimulated further research and has raised new, unforeseen problems. In short, evaluate the importance of his work in terms of its breadth of applicability. DO NOT consider other aspects of his performance--ONLY the creativity of his work.

Hypothetical breakdown of how a sample of 100 scientists might be distributed across the descriptive statements.

1\.

2. — 5% His work has demonstrated very little creativity or originality. It usually has provided no more than a rather simple solution to the immediate problem.

3\.

4. — 10% His work has generally been a result of standardized approaches demonstrating little creativity. At its best, it has produced minor innovations which have been limited to the immediate problem.

5\.

6. — 20% He has been moderately successful at solving problems and eliminating difficulties that arise in his work. Generally, his work has been of moderate importance but definitely narrow in applicability.

7\.

8. — 30% His work has occasionally demonstrated some relatively original approaches and new ideas in the solution of some difficult problems but it has been mostly familiar or conventional and usually somewhat narrow in terms of its implications and applicability.

9\.

10. — 20% He has more than the average number of new ideas, and his work has often been fresh and original. He has pointed out ways that techniques or results could be used beyond their original purpose.

11\.

12. — 10% He has conclusively demonstrated a high degree of creativity. His work has presented comprehensive solutions to difficult problems, with some significant implications applicable to other areas of research.

13\.

14. — 5% The impact of his work has been quite exceptional. His very creative solutions to very complex problems have broad generality and have even opened up important new areas of investigation with wide implications.

15\.

from the scale which best describes each man in the parenthesis beside his name.

Consider the man's overall work performance, his total contributions as well as his specific assignments, whether more observable (productive in reports, patents, etc.) or less observable (hard worker, makes useful suggestions, stimulates thinking of others, etc.). In terms of traits, consider his accuracy, thoroughness, initiative, originality, attitude, cooperation, the degree to which he gets along with others, etc. In other words, consider each man's total contributions in comparison with his associates and with the other scientists you have known.

( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )
( )

Hypothetical breakdown of how a sample of 100 scientists might be distributed across the descriptive statements.

5%
10%
20%
30%
20%
10%
5%

1.
2. — INFERIOR
3.
4. — NOT IMPRESSIVE OR POOR
5.
6. — SOMEWHAT BELOW AVERAGE
7.
8. — ABOUT AVERAGE
9.
10. — SOMEWHAT ABOVE AVERAGE
11.
12. — SUPERIOR
13.
14. — TRULY OUTSTANDING
15.

(WHEN YOU HAVE RATED EACH MAN, TURN THE PAGE.)

Name of Ra.

____

Names of Men to be Rated	Write the number from the scale which best describes each man in the parenthesis beside his name.
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
______________________	( )
	( )

The scale below is for the purpose of obtaining your estimate of how confident you are that your previous ratings for each man represent an accurate appraisal of his performance. On the basis of how long you have known each man and how closely you have worked with him, indicate how confident you are that your four ratings of each man represent an accurate evaluation of his performance. Write the number of the statement which best describes your confidence estimate for your ratings within the parentheses by his name.

1. VERY CONFIDENT
2. QUITE CONFIDENT
3. SOMEWHAT CONFIDENT
4. NOT AT ALL CONFIDENT

Your cooperation and effort in completing these forms is greatly appreciated.